# OUTCOMES AND EFFECTIVENESS OF CHILDREN'S HELPLINES

## A SYSTEMATIC EVIDENCE MAPPING

Mariya Stoilova, Sonia Livingstone, Sheila Donovan

September 2019

NSPCC

LSE Consulting

LSE Media and Communications

# Impact and Evidence series

This report is part of the NSPCC's Impact and Evidence series, which presents the findings of the society's research into its services and interventions. Many of the reports are produced by the NSPCC's Evidence (formerly Evaluation) department, but some are written by other organisations commissioned by the society to carry out research on its behalf. The aim of the series is to contribute to the evidence base of what works in preventing cruelty to children and in reducing the harm it causes when abuse does happen.


Photography by Ross Bolger. The child pictured is a model.

# Contents

# Tables and figures

# FOREWORD

NSPCC are delighted to publish this evidence review of child helplines, conducted by LSE Consulting on our behalf. Being an evidence-informed organisation is a key priority for NSPCC. We seek to learn about, and share, what we know works well to safeguard children and prevent abuse, whether about our own services or the wider children and young people's sector.

NSPCC's organisational mission since 2016 has focused on achieving five goals to make five million children safer by 2021. Our helplines play a key role in that mission, providing the means for children to speak, and adults to take action, about abuse. Childline uses feedback from children and professionals to learn about when and how it works well, and continually develops in order to best meet the needs of those who use it.

There are several methodological and practical challenges to carrying out research and evaluation in a confidential helpline environment, so we wanted to learn about further ways to do this effectively drawing on the experience of others who have done so. We also wanted a comprehensive and current picture on the outcomes of child helplines, in order to enhance our own internal understanding of effectiveness.

Whilst being of enormous value to ourselves, we also hope that bringing together this evidence in one place will be of use to anyone else involved in delivering or evaluating child, and indeed other, helpline services. We would like to extend our thanks to the authors- Mariya Stoilova, Sonia Livingstone and Sheila Donovan- for producing such a rich and clear picture of the types of child helpline outcomes captured in the research literature to date.

Kate Stanley
Director Strategy, Policy and Evidence

# 1. Overview of key findings

The work of child helplines is vital for reducing children's risk of harm and vulnerability and improving child protection and wellbeing. Starting in 1979 with the first telephone helpline for children – the Dutch De Kindertelefoon – child helpline services have spread across the world and are estimated at present to take nine million calls annually (Dinh et al, 2016). Child helplines are now recognised as a key component of child protection services that contribute to creating accessible and child-friendly reporting systems and can help to ensure the implementation of child rights (UNICEF, 2007).

Understanding the evidence regarding the effectiveness of child helplines can inform improvements to helplines and related child protection mechanisms. This review maps the existing literature on the effectiveness of child[1] helplines in order to identify outcomes and approaches to measuring their effectiveness. Comprising an extensive search of 18 databases, the review is based on a systematic mapping of the evidence (Grant & Booth, 2009; Gough, Thomas and Oliver, 2012).

## 1.1 Identifying helpline outcomes

The systematic evidence mapping identified a range of outcomes associated with the work of helplines, ranging from the individual level outcomes (such as changes to attitudes, knowledge and behaviour) to social (or societal) level outcomes (notably, changes in relations between organisations, community dynamics, social attitudes and protection policy). Most evidence is focused on outcomes related to the individual service user, with few evaluations of interventions designed to affect the organisational, community or societal level. This is partly because helplines themselves generally aim to benefit individual users first and foremost, but also because of the methodological difficulties of attributing effects beyond the individual level (Coveney et al, 2012). Most evidence relates to external support offered by helplines (referrals, medical help, brokering with other agencies), satisfaction with the service, and changes in feelings and attitudes. Generally, outcome selection is not based on particular outcome models, theoretical approaches or even organisational aims but is, rather, pragmatic, tending to be based on outcomes most amenable to measurement. However, this report highlights good models for the systematic evaluation of helpline outcomes.

1 See Appendix 1 for key definitions, including 'child', 'helpline', 'outcomes', 'impact', some methodological terms, etc.

## 1.2 Key positive outcomes include a competent, efficient and approachable service

The evidence demonstrates that helplines can achieve improvements in relation to a range of child-centred outcomes, notably children's wellbeing, self-confidence, levels of anxiety and distress, and ability to deal with their current situation. It shows that children value professionals they can trust, who are effective, knowledgeable and available to them (Cossar et al, 2013). Benefits sought and valued by children include the competence of the service provider, feeling welcome, and being able to talk about issues of concern. In some cases, help-seekers put more emphasis on emotional support than on problem-solving (Law et al, 2015). Indeed, the most positive effect of a helpline for a child can be establishing in the child's mind that they are being listened to and understood (Coveney et al, 2012). However, most evaluations concern the effects of a service as a whole, with scarce evidence regarding which elements of the service are responsible for the positive outcomes. This makes it difficult to draw transferable conclusions about helpline effectiveness, or about why we see more improvement on some outcomes than others.

## 1.3 Limitations to the positive outcomes

While the overall effects of child helplines are generally positive, helpline effectiveness varies depending on the outcomes measured. Fewer effects are observable in relation to skill development or changes in individual, family and social behaviour and practice. This may be due to the nature of the service offered, or to the methodological challenges in capturing these effects. In addition, some evaluations discovered no effect or even negative effects on some service users. The children who are worse off after using the service are usually those who are unable to achieve the outcomes they expect, who are unhappy with the advice and help offered or who find that, while the advice is good, it is difficult to follow. *Setting realistic and measurable helpline outcomes that capture the service performance can help to identify gaps and improve support, particularly for the children who experience the least benefits.*

## 1.4 Extended helpline support in the digital era

The majority of the helplines discussed in the evidence review offer additional internet- or mobile-mediated elements of the service, such as chat or email services, forums, text messages or apps. There is a growing number of studies on the effectiveness of online services (Fukkink & Hermanns, 2009; Dinh et al, 2016; Cronin et al, 2017), possibly facilitated by the comparative ease of gathering pre-session information (e.g., via online registration in order to use the service). The existing evidence suggests that online support offers

distinct benefits over telephone services. The positives include more efficiently engaging with children seeking help, particularly seldom heard groups, such as lesbian, gay, bisexual, transgender and queer (LGBTQ) youth, or those struggling with mental health issues or with a speech impairment (Burns & Birrell, 2014; Childline, 2016). Online services are usually seen as more accessible (for those with internet access) as they have no geographical or time boundaries. Arguably, for some they can offer greater emotional safety and security due to the reduced emotional proximity to the counsellor. Compared with phone calls, online chat with a counsellor can provide greater privacy and anonymity, as well as a better opportunity to plan the interaction (Bambling et al, 2008; Andersson & Osvaldsson, 2011; Cossar et al, 2013; Haner & Pepler, 2017). This makes online support preferred among children who feel uncomfortable contacting a telephone helpline, particularly those with more complex and emotionally charged issues (King et al, 2006; Law et al, 2015; Haner & Pepler, 2017). However, particular challenges arise from online communication, including the difficulty of recognising distress due to the lack of non-verbal and paralinguistic information (expression, voice, intonation, pauses), resulting in the need to employ compensatory strategies (e.g., emoticons, expressing emotions in text, discussing misunderstanding if it occurs) (Bambling et al, 2008; Fukkink & Hermanns, 2009; Sindahl, 2013). Additional challenges posed by online services themselves, such as inequality of access and skills, and the trend towards technologically driven service provision suggests more attention is needed in this area in research, policy and practice, with heightened acknowledgment of the prevailing inequalities.

## 1.5 Effectiveness of the different channels

While it is common for helpline services to employ different channels for support, the review found only few studies that compared their effectiveness – mostly telephone and online support (King et al, 2006; Fukkink & Hermanns, 2009; Law et al, 2015) – and one study (van Dolen & Sindahl, 2017b) comparing more channels (including email and text messaging). The types of outcomes measured vary across the studies and they paint a rather mixed picture of channel effectiveness, which makes it hard to draw robust conclusions. Some studies found online support more effective overall (Fukkink & Hermanns, 2009; van Dolen & Sindahl, 2017b) or in relation to particular outcomes, such as improved wellbeing, decreasing the perceived burden of the problem and increased perceived clarity of the problem (Fukkink & Hermanns, 2009; Law et al, 2015). Yet other studies found that telephone support achieves greater improvements in relation to outcomes like hopefulness, self-efficacy, distress (Law et al, 2015) or is overall more effective than online services (King et al, 2006). Such

disparities might be explained by differences in the service users or the service itself, as well as with how outcomes are identified and measured. *Future evaluation efforts should concentrate on exploring the efficiency of the different channels and how each can be used most effectively to support the different needs or groups of users.*

## 1.6 Evidence-driven service

The evidence mapping[2] found that helplines generally strive to monitor their performance and use evidence to improve their support to children. A wide range of evidence-gathering and evaluation methods exists, from small-scale micro studies to more comprehensive multi-method designs, and from single-point 'snapshots' to comparative design studies with a longitudinal focus. While some studies rely on theoretically and empirically informed designs and map the organisational aims onto the service outcomes, others have a patchier approach and do not always draw on clearly defined outcomes in their assessments.

The helpline sector faces substantial challenges in evaluating its effectiveness. This is because of both the struggles of individual helplines (often small organisations) to carry out high-quality evaluation methods, and the lack of established and recognised models and outcome measures within the sector. Effectiveness is measured by collecting different types of evidence related to the work of the helpline. The different approaches can be grouped thus: effectiveness models (clearly identified aims, outcomes, robust evaluations based on effectiveness criteria and clear pathways to effectiveness); measuring outcomes (using evaluations to gather empirical evidence on improvement in relation to clearly identified outcomes); studies of the process of intervention (demonstrating good practice during the intervention process drawing on theoretical models or prior empirical evidence of what works well); evidence on outputs, activities and beneficiaries affected; or a combined approach (blending more than one of the above approaches). Half of the helpline evaluations in the review were sufficiently robust to demonstrate effectiveness of the service (via quasi-experimental, ex post facto or experimental design) while the other half could only demonstrate potential effectiveness (based on exploratory, case study, observational or ethnographic, narrative or discursive or interview/focus group design).

2 See Section 3.4 and Appendix 2.

## 1.7 Voicing children's concerns

The main strength of the evidence base is that it represents children's concerns and issues in a comprehensive way. Most child helpline evaluations provide a good overview of the key issues children face, often drawing on case studies that vividly document children's experiences. Some helplines particularly prioritise child representation, advocacy and rights and others provide 'on-demand' child support and evidence in response to a particular socially relevant issue. More efforts to provide evidence of the issues of concern to children themselves could open up the decision-making process and ensure children's issues are represented at policy level.

## 1.8 Circulation of good practice

Helplines do not work in isolation and effective collaboration within the sector – with referral agencies, educational and governmental institutions and the wider public – is often mentioned as key to effective child protection. Still, there is little evidence related to effects occurring at the level of family, community, cross-agency networks or society as a whole. Given substantial efforts to create change at all these levels, these evidence gaps are most likely the result of methodological difficulties to attribute such effects. With so many other factors at play, establishing exclusive and direct causality between a helpline's activities and a change in a child's life, community or society can be difficult (Child Helpline International and Oak Foundation, 2017: p.26). This identifies a need for more comprehensive mapping of social-level outcomes, which are measurable and correspond to the areas of work of child helplines. This may require collaborative evaluations, along with tracking longer-term impact and even establishing evaluation standards across the sector. Some good examples here are the efforts of Child Helpline International and Insafe to establish international evaluation models and standards to be used by the helplines in their networks (Dinh et al, 2016; Child Helpline International and Oak Foundation, 2017), or evaluations that compare the effectiveness of different actors in relation to a particular issue, such as child maltreatment or sexual exploitation and abuse (Child Helpline International and UNICEF, 2017; Fry et al, 2017).

## 1.9 Gaps and challenges

- ***Brief and confidential contact**:* the effectiveness of helplines is difficult to establish in a context where the contact with callers may be brief, singular and anonymous. In many cases (particularly related to vulnerable groups or acute situations), there is little scope for caller feedback or follow-up evaluations (Coveney et al, 2012). While helplines have different approaches based on how they see their role (providing information and signposting or protecting children through involving other agencies), confidentiality poses challenges to measuring outcomes and following up with service users.
- ***Sensitivity of issues**:* in many cases, children contact the helpline to discuss sensitive issues and may have difficulties articulating the problems they face or coming to terms with their effects. This makes evaluations particularly challenging, both ethically and in practice, with many children refusing to participate.
- ***Diversity of factors**:* another difficulty in establishing the effectiveness of the helpline, especially in relation to longer-term effects, lies in the diversity of factors that may influence the outcomes, including behavioural, psychological and socioeconomic, use of other services, etc.
- ***Preventative effects**:* some of the helpline work relates to prevention of negative effects, which is particularly hard to measure and demonstrate. Looking at population data and being able to demonstrate change in the social trends (e.g., in the incidence of child sexual abuse, violence, bullying, suicide, etc.) is a possibility, but as a relatively small proportion of the population contacts helplines, it is unlikely that preventative effects can be convincingly measured with population data. Furthermore, for helplines that offer support in relation to a broad spectrum of issues, seeking impact at a social level would be even more challenging.
- ***Long-term effects**:* there is relatively little empirical evidence that positive changes at the end of the intervention would lead to long-term positive outcomes, and evaluations with longitudinal elements are particularly rare in the sample of the review. For the few studies that attempt to trace longer-term effects (Andersson & Osvaldsson, 2011; Fry et al, 2017), another challenge lies in the difficulty of demonstrating with certainty that the observed benefits are the result of the service use rather than the spontaneous improvement or use of other services. The way to ascertain this effect is to conduct a longitudinal investigation with a control group, but none of the studies reviewed used such a design. In fact, such a design is probably impractical for ethical, methodological or cost reasons.

- ***Child-friendly methods**:* better outcomes data often means more intrusive research methods, which may be unsuitable for children – especially those who have self-identified as needing helpline support. The evaluation of child helplines needs to acknowledge the primacy of children's wellbeing and the need for child-friendly methods (such as evaluations using participatory methods or more visual methods for communicating with younger children or those with limited literacy or capacity to concentrate). While these methods do not necessarily follow recognised 'gold standards' for outcome evaluations, and are underrepresented in the evidence, they are important for reaching out to children in ways that are comprehensible to them and respect children's needs and rights.

## 1.10 Overcoming barriers to effectiveness

- ***Multiple measures and indicators***: one of the strengths of the evaluations included in the review relates to the use of mixed methods to demonstrate the effectiveness of the service. Given the sensitive settings in which helplines operate, the number of issues raised and the various channels of support that are used at present, combining a range of methods, including both quantitative and qualitative measures, can provide stronger evidence.
- ***Quality of evidence**:* it might be beneficial to distinguish between 'hard' and 'soft' outcomes and to use different types of evidence to demonstrate effectiveness. Using a combination of some less robust quality data that a helpline can gather longitudinally and a single high-quality evaluation can offer an effective approach to demonstrating effectiveness.
- ***Using standardised measures**:* the use of standardised measures with established reliability, validity and applicability to children is a good way of producing robust evidence of effectiveness.
- ***Mapping aims, outcomes and measures**:* having a clear model of how the organisational aims match the desired outcomes and how the evidence can support the demonstration of effectiveness is crucial.

# 2. Acknowledgements

We would like to thank the team of the National Society for the Prevention of Cruelty to Children (NSPCC) – Richard Cotmore, Chloe Gill and Tamar Moselle – for their support in shaping and guiding this evidence review. We are also grateful to Rebecca Harrison, Karina Kinsella, Ellie McCoy, Jane Oyston, and Dr Zara Quigg (Public Health Institute, Liverpool John Moores University) for sharing their early findings from the Childline stakeholder engagement.

We would also like to thank the experts consulted for their valuable insights and suggestions: Magdalena Aguilar and Helen Mason (Child Helpline International), John Carr (CHIS), Anne Collier (The Net Safety Collaborative), Julia Davidson (Middlesex University), Mikael Drackner (BRIS), Ged Flynn (Papyrus), Liam Hackett (Ditch the Label), Laura Higgins (South West Grid for Learning), Rebecca Hitchen (Rape Crisis), Peter Irgens (Västerbottens läns landsting), David Marjoribanks (Relate), Hans Martens and Karl Hopwood (Insafe), Neil Melhuish and Edgar Pacheco (NetSafe New Zealand), David Miles (British Board of Film Classification), Brian O'Neill (Dublin Institute of Technology), Elizabeth Scowcroft and Stephanie Stace (Samaritans), June Statham (Institute of Education, UCL), Trine Natasja Sindahl (Børns Vilkår) and Willemijn van Dolen (University of Amsterdam).

We would also like to express our gratitude to Heather Dawson (British Library of Political and Economic Science) for her expert suggestions and guidance in selecting the most relevant databases.

# 3. Introduction

## 3.1 Aims of the evidence mapping

Understanding the evidence regarding the effectiveness of child helplines can inform improvements to helplines and related child protection mechanisms. The aim of the review is to map systematically the existing literature on the effectiveness of child helplines. The focus is on the definition of outcomes, identifying the range of approaches taken to measuring effectiveness, and evaluating the strength of the evidence. This analysis underpins the recommendations to the NSPCC developed in the last section of this review regarding the future evaluation of Childline.

## 3.2 Context

The work of child helplines is vital for reducing children's risk of harm and vulnerability and improving child protection and wellbeing. Starting in 1979 with the first telephone helpline for children – the Dutch De Kindertelefoon – child helpline services have spread across the world and are estimated at present to take nine million calls annually (Dinh et al, 2016). Child helplines are now recognised as a key component of child protection services that contribute to creating accessible and child-friendly reporting systems and can help to ensure the implementation of child rights (UNICEF, 2007).

In England, it is estimated that there were 389,430 children 'in need' in 2017 (children deemed unlikely to achieve reasonable health and development without the provision of services), which equates to over 33 children per 1,000 aged under 18. Both the number of looked-after children and children subject to child protection plans were on the rise in 2017 – there were 72,670 looked-after children, with nearly two-thirds of these cases being due to abuse or neglect (DfE, 2017b), and over 51,000 children were subject to child protection plans during the same year (DfE, 2017a). Of the available child helplines, Childline, which is one of the most well-known telephone counselling services, has helped over 4 million UK children since it was launched in 1986 (NSPCC, 2016). In 2016/17, it provided 295,202 counselling sessions to children (Rantzen & NSPCC, 2017).

The work of the child helplines over the past decades has been influenced by two important developments. First, there has been an increased integration of digital technologies into children's everyday lives, communication and service provision that has created a high and potentially rising demand for online support for children. Seventy-one per cent of the NSPCC's counselling sessions now take place online (Rantzen & NSPCC, 2017). In March 2017, the NSPCC introduced

a Childline app (*For Me*), the first free-to-download app to provide direct counselling for children through a mobile device (Rantzen & NSPCC, 2017), to respond to the need for digitally mediated support.

Second, the work of the helplines has been influenced by increased recognition of the importance of evidence-based support and the necessity of demonstrating service effectiveness. Following the Children Act 2004, an Integrated Inspection Framework was created to assess the extent to which the new Children's Services Authorities contribute to better outcomes for children, with key areas identified as 'being healthy', 'staying safe', 'enjoying and achieving', 'making a positive contribution' and 'achieving economic wellbeing' (Ofsted, 2004; Hudson, 2005). While this framework, among other efforts, emphasises the importance of using evidence to improve available services and ensuring key outcomes are met, this remains an under-researched field, not least because of severe methodological challenges, as this review will show.

## 3.3 NSPCC

The National Society for the Prevention of Cruelty to Children (NSPCC, initially the Society for the Prevention of Cruelty to Children) was founded in 1884 with the aim of protecting children and supporting vulnerable families (NSPCC, 2008). Since it was established, it has grown and expanded, encompassing 180 teams and services helping to protect children across England, Wales, Northern Ireland and the Channel Islands. In 2006, Childline formally became part of the NSPCC (NSPCC, 2008). Childline was launched in 1986 as a confidential telephone service to support children in difficulty. Children contact the helpline to seek help in relation to a wide range of issues, including abuse, bullying, self-harm, eating disorders, depression and other aspects of mental health (NSPCC, 2016). To respond to the changing needs of children, Childline has grown and extended its services to include digital and online platforms, and now also offers support via online chats, email, an app and an online forum for peer support. With the help of around 1,400 counsellors who volunteer in 12 bases around the UK, Childline now offers 24-hour support throughout the year (NSPCC, 2016).

In addition to Childline, the NSPCC runs an adult helpline (phone and online service) supporting parents and professionals, as well as a number of dedicated helplines that work in partnership with other agencies and offer advice on issues like radicalisation, female genital mutilation (FGM), online safety, gangs and modern slavery. The NSPCC also works directly with children and families at 29 service centres across the country.[3]

3 More information about the NSPCC's services is available at www.nspcc.org.uk/services-and-resources/

The NSPCC has an evidence-based approach to service development and is carrying out over 25 separate evaluations of its services for children and families.[4] This particular review was commissioned in order to inform the future planning of an impact study into the effectiveness of Childline. While there is increasing understanding of the benefits Childline offers its service users, there are still gaps in understanding the outcomes for children who contact Childline. This review aims to assist in establishing a solid evidence base against which the effectiveness of Childline can be evaluated, and supplements another study that further develops and strengthens the Childline Theory of Change[5] by engaging key stakeholders in a series of workshops with children, Childline counsellors and their managers. The evidence mapping also updates the findings from an earlier evidence review carried out by the NSPCC in 2008, which found that there has been little research into the effectiveness of telephone counselling (Cotmore & Fernandes, 2008).

## 3.4 Methodology

The review systematically maps the existing literature on the effectiveness of child helplines with a particular focus on identifying helpline outcomes and approaches to measuring effectiveness. The research questions guiding the review are:

- How do children's helplines define service 'outcomes' and are there any gaps?
- What is the established criteria for 'effectiveness' of services, and what approaches are used to measure the effectiveness of children's helplines?
- What is the strength of the existing evidence base for the effectiveness of children's helplines, and what recommendations can be made about best approaches?

4 For more on the evidence of the impact of NSPCC services, see www.nspcc.org.uk/services-and-resources/impact-evidence-evaluation-child-protection/

5 'Theory of Change is an outcomes-based approach that applies critical thinking to the design, implementation and evaluation of initiatives and programmes intended to support change in their contexts.' (Vogel, 2012: p.3)

Table 1: Coded results by category

| Category | Type | Number of studies |
|---|---|---|
| Type of study (*n*=73) | Primary research | 51 |
| | Secondary data analysis | 11 |
| | Conceptual | 6 |
| | Systematic review or literature review | 5 |
| Type of data (primary research, *n*=51) | Qualitative | 19 |
| | Quantitative | 17 |
| | Mixed method | 15 |
| Methodology (primary research) | Experimental | 1 |
| | Quasi-experimental | 24 |
| | Exploratory | 5 |
| | Case study | 2 |
| | Ethnographic or observational | 7 |
| | Narrative or discursive | 6 |
| | Interview/focus group | 6 |

The review is based on a systematic mapping of the evidence (Gough et al, 2012; EPPI-Centre, 2018) comprising a more comprehensive search strategy that allowed for the inclusion of a broader range of sources, such as end-of-year reports, policy recommendations, methodological guides and case studies. The systematic mapping included an extensive search of 18 databases (see Appendix 2, Table 5) resulting in the identification of 1,421 search results and a two-phase process of screening of results based on predefined inclusion criteria. The screening reduced the results to a final sample of 73 studies, which were analysed using a coding frame constructed to meet the requirements of the review. To ensure maximum insight and rigour, the team also requested input from a range of experts (see Acknowledgements) on recommended research sources that were included in the search sample. For a detailed overview of the methodology including search terms, databases searched, inclusion criteria and screening and coding procedures, see Appendix 2.

## 3.5 Structure of the review

This review begins with an overview of key findings about the current developments of child helplines and then moves on to discuss the findings related to helpline outcomes (models, types of outcomes) and effectiveness (approaches to effectiveness, evaluation designs and barriers to effectiveness). A set of case studies is included as a way of highlighting good practice or relevant evidence-gathering suggestions. The review concludes with recommendations. A detailed methodology, coding framework and list of the reviewed studies are available in the Appendices. The coded studies are available as a separate review supplement.

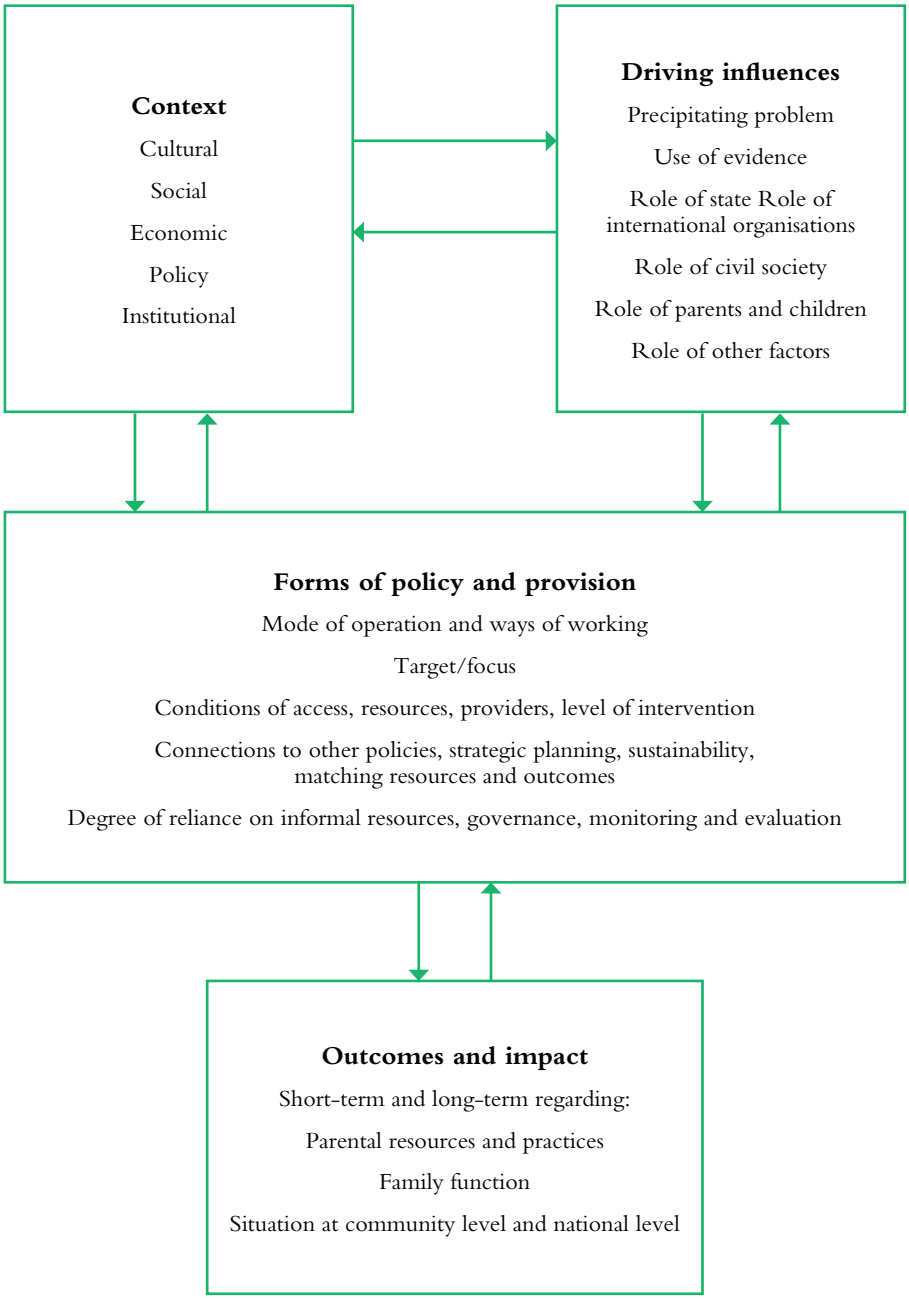
# 4. Helpline outcomes

The sample of evaluations identifies a range of outcomes associated with the work of helplines spanning the individual to the social level. Still, the majority focus on the outcomes associated with the service user and much less attention is dedicated to identifying any effects occurring at the organisational (e.g., change of practice), community (e.g., prevention efforts) or social level (e.g., attitudes towards child protection). This can be explained both with the focus and aim of the helplines and the methodological difficulties of attributing effects beyond the individual user (Coveney et al, 2012). In most cases, the selection of particular outcomes is not explicitly associated with any outcome models, theoretical approaches or organisational aims, but most studies suggested that outcome performance would contribute to the planning of future activities. In the few exceptions where the selection of outcomes was justified, the studies referred to Theory of Change or organisational logic models (see Shelter, 2011; Daly et al, 2015; Law et al, 2015; Child Helpline International and Oak Foundation, 2017) and bottom-up development of outcome frameworks (Daly et al, 2015) as methods for identifying service outcomes.

## 4.1 Outcome models

The review found a small number of comprehensive models that map the terrain of intervention outcomes (see Daly et al, 2015; Law et al, 2015; Child Helpline International and Oak Foundation, 2017). They all offer a combination of two factors: location where the outcomes occur (from individual to social level); and time when these outcomes occur (immediate/short-term, intermediate and final/long-term outcomes). One of these models, developed by Daly and colleagues (2015) in relation to parental support services, puts the analysis of outcomes and impact in a wider framework that takes into consideration the context (cultural, social, economic, policy, institutional), the driving influences (including the role of the state, civil society, international agencies, parents, use of evidence etc.) and forms of policy and provision (such as conditions for access, resources, providers, governance, monitoring and evaluation, strategic planning) (see Figure 1). Daly and colleagues acknowledged the need to identify the negative and unintended outcomes arising from cases where the intervention did not achieve the positive outcome planned.

Figure 1: A comprehensive framework for analysis of support




Source: Daly et al (2015)

This model can be useful for conceptualising the terrain in which helplines work, identifying how the intervention outcomes might be fostered or impeded by external dynamics, and planning better the pathways to effectiveness. The framework developed by Daly and colleagues (2015) is also useful in identifying the gaps in the existing evaluations of child helplines. As this review demonstrates, a majority of the evaluations we sampled engage predominantly with short-term outcomes and those related to the individual child. Outcomes related to the wider social network of the service user are much less frequent, and when they are taken into consideration, it is mostly by helplines that provide additional face-to-face services. Discussion of the situation at community or national level is also rarely present, but a number of good examples exist, as we discuss later. This observation corresponds to the findings of a recent systematic review of social and behaviour change interventions addressing violence against children, which also discovered much fewer interventions focused on engagement at the organisational or community level (Cronin et al, 2017).

Drawing on the existing models of child helpline outcomes that the systematic mapping identified (see Daly et al, 2015; Law et al, 2015; Child Helpline International and Oak Foundation, 2017), and supplementing them with findings from other evaluations included in the review, we propose a helpline outcomes framework (see Table 2 below), which records in a comprehensive way the range of outcomes discussed in the sample of evaluations. The framework is then used to systematically organise the findings and examples of evaluation of each type of outcome discussed, noting where discrepancies in the amount of existing evidence exist. Overall, the further down the table, the less evidence and evaluations we found in our sample.

The framework summarises the outcomes from the point of a helpline (as opposed to other forms of intervention, such as face-to-face child support), and focuses mainly on child users due to the scope of the systematic mapping. However, some of the outcomes identified might be relevant to adults, judging from the small number of adult-related helplines we reviewed (for parents or educators supporting children, or perpetrators of child sexual exploitation).

## Table 2: A framework for analysis of child helpline outcomes

| Level | Type of outcome | Outcome details<br>↑ **increase/improvement sought for effectiveness**<br>↓ **decrease/reduction** |
|---|---|---|
| Individual level | External support offered | ↑ Child referred to other services (e.g. medical help) |
| | | ↑ Follow-up calls or contact, where applicable |
| | | ↑ Well-balanced referral response (not to overwhelm the child) |
| | | ↑ Brokering with other agencies |
| | Satisfaction with the service | ↑ Feeling heard, understood, taken seriously, supported, at ease |
| | | ↑ Likely to recommend the service |
| | | ↑ Repeated use of service, preference over other services |
| | | ↑ Received useful information, support |
| | | ↑ Strong therapeutic alliance |
| | Changes in feelings and attitudes | ↑ Hope, confidence, belief in oneself, self-efficacy |
| | | ↑ Safety, security, trust |
| | | ↓ Distress, isolation, stress, level of perceived burden |
| | Changes in awareness and knowledge | ↑ Awareness of personal strengths and resources |
| | | ↑ Awareness of choices and options |
| | | ↑ Formulation of a plan of action by the child |
| | | ↑ Mental and emotional health literacy |
| | | ↑ Knowledge about issues affecting their life |
| | | ↑ Understanding own preferences and values |
| | | ↑ Knowledge of community resources |
| | Skills development | ↑ Coping, self-care skills, independence |
| | | ↑ Problem-solving skills |
| | | ↑ Ability to articulate own wellbeing needs |
| | Changes in individual behaviour and practice | ↑ Intentional use of personal strengths, capacities and skills |
| | | ↑ Practising effective coping and self-care strategies |
| | | ↑ Additional help-seeking, as required |
| | | ↑ Use of community resources, as required |
| | | ↑ Engagement in social networks, community, education |
| | | ↑ School attendance and performance |
| | | ↓ Behavioural problems |
| Interpersonal level | Changes in family and social behaviour and practice | ↑ Strengthened relationships and increased support (family, peers, social networks) |
| | | ↓ Child exposure to risks |
| | | ↓ Anti-social or hurtful behaviour |

| Level | Type of outcome | Outcome details<br>↑ increase/improvement sought for effectiveness<br>↓ decrease/reduction |
|---|---|---|
| Organisational level | Changes in organisational practices and capabilities | ↑ Robust screening, management and training and capacity building of staff and volunteers |
| | | ↑ Funding support and fundraising capability |
| | | ↑ Social support for the organisation |
| | | ↑ Contribution of volunteers |
| | | ↑ Toll-free status of the helpline, increased capacity and diversity of services, a range of available resources |
| | | ↑ Adequate support to children by the helpline |
| | | ↑ Self-assessment of impact by staff |
| | | ↑ Producing timely data |
| Cross-organisational and community level | Changes in inter-organisational practices and community dynamics | ↑ Extensive referral networks, agreements and established protocols with reporting hotlines, law enforcement and other referral partners |
| | | ↑ Collaboration and coordination among agencies within the sector |
| | | ↑ Strengthening partnerships and campaigns |
| | | ↑ Quality input from stakeholders |
| | | ↑ Accessibility of children to support and referral agencies |
| | | ↑ Children's issues are voiced, advocacy |
| | | ↑ Recognition/assessment of effectiveness by experts, key informants, stakeholders |
| | | ↓ Rates of placement of children in alternative care |
| | | ↓ Rates of family conflicts, violence and abuse |
| National and international level | Changes in policy, protection and society | ↑ Protection of children's rights |
| | | ↑ Placing children at the centre of decision-making |
| | | ↑ Promoting wellbeing of children |
| | | ↑ Change in social values and practices |
| | | ↑ Early prevention and reduced financial cost |
| | | ↑ Awareness of risks to children and protective mechanisms |
| | | ↑ Portfolio of policies and interventions based on evidence |
| | | ↑ Standards of delivery, strengthened accountability and governance structures |
| | | ↑ Well-trained and active sectoral force |
| | | ↑ International collaboration and knowledge exchange |

Sources: Mental Helplines Partnership (2003); King et al (2006); Daly et al (2015); Law et al (2015); Dinh et al (2016); University of Central Lancashire and National Children's Bureau Research Centre (2016); Child Helpline International and Oak Foundation (2017); Child Helpline International and UNICEF (2017); Cullen et al (2017); van Dolen & Sindahl (2017a)

### 4.1.1 External support

This first type of outcome refers to some action taken by the helpline by cause of the contact that results in a tangible and practical support offered outside the initial interaction. This might include referrals made to other agencies, brokering access to additional services (e.g., treatment facilities, shelters, rehabilitation centres) or acting on behalf of the caller (e.g., in front of council housing agencies or the prison service) (Calderón et al, 2017). Some of the helplines represented in the sample also offer follow-up calls and report on the number of calls, some evaluating their effectiveness (see, for example, Donaldson & Irwin, 2017; Goldsmith, 2017). The extent to which child helplines engage in the provision of external, more practical support varies significantly based on the approach of the helpline itself but also on the availability of other services to which users can be signposted. The case study below, on Childline India, demonstrates a more intensive approach in a context where some of the practical support is carried out by the helpline itself rather than being referred to other services (such as the police or social services).

On the other hand, using referrals seems a universal approach to demonstrating service outcomes. Most of the studies in the sample discuss the number of referrals made to other agencies, sometimes supplementing this data with information on an increase in the number of referrals in comparison to previous years, referrals to different agencies or by primary concern leading to the referral (Childline, 2016). Referrals can also be applied as a basis for cross-country comparisons; for example, Fry and colleagues (2017) used secondary data on child maltreatment in Australia, Canada, England and the US, and compared the number of referrals per 1,000 children to estimate the prevalence of the issue in the different contexts.

The suitability of referrals as key outcomes measures, however, is not without criticism – the lack of follow-ups to establish if the referral resulted in positive outcomes is pointed out, as well as the inability of this outcome measure to capture the full service of the helpline, for example, in cases where the support was offered 'in house' and a referral was avoided (Statham & Carlisle, 2004; Mishara et al, 2007; Gallagher, 2013).

### Case study: Childline India

Established in 1996, Childline India includes a helpline for adults and children in distress available in 366 districts in 34 state areas in India (as of 2015) (Childline India, 2018). Up to 2015, the helpline had received 36 million calls. Two studies in the sample discussed the work of the helpline – Dottridge (2008), who used an ethnographic method, and Fernandes (2006), who used a mixed-methods approach employing site observations as well as focus group discussions and interviews with approximately 300 children (street children, children in residential homes and children living in slum communities), and questionnaires and interviews with 40 frontline workers of the child helpline.

Both studies emphasise the substantial practical support that Childline India offers, often based on the contribution of child volunteers who are involved in spreading the word about the helpline on the street and in communities, calling on behalf of others who need help, and assisting with the identification of trafficked children by monitoring railway stations (Fernandes, 2006). Such volunteer contribution allows the helpline to offer more practical support and have a more interventionist approach. Childline India has taken on a brokering role in encouraging and facilitating collaboration between the different child protection agencies and creating a referral system, bridging the federal, district and local levels.

### 4.1.2 Satisfaction with the service

Arguably, the most prevalent approach to service evaluation is measuring user satisfaction with the service. This includes gathering information on overall satisfaction as well as on a number of particular service-related issues, such as helpfulness of the service or information, intention to act on advice received, fulfilment of expectations, likelihood of recommending the service or the ease of navigation of the service elements (particularly in relation to online platforms). For example, to determine whether users of Kids Help Phone (Canada) are satisfied with their telephone and counselling experience, Law and colleagues (2015) asked questions about children's expectations, whether they would use the helpline again and whether they would recommend the service to a friend. A number of evaluations also asked the users about access to and satisfaction with other services and how the helpline compared to them.

As part of service satisfaction, some studies also focus on the nature of communication with the counsellor. The clarity of the information received was shown to be an important factor for an effective service by a number of studies (Downing & Cook, 2006; Reubsaet et al, 2006; Hutson & Cowie, 2007; Bambling et al, 2008; Fukkink & Hermanns, 2009; Sharratt et al, 2014). Others point to the significance of a welcoming and friendly interaction; non-judgemental and respectful treatment; and a kind, sympathetic and understanding response. User satisfaction and positive outcomes depend on factors like the expression of empathy, validation of emotions, giving moral support, reframing, talking about own experience and offering a follow-up call (Mishara et al, 2007). A knowledgeable and professional way of handling the situation is related to user satisfaction, but also to whether a user reports that they will act on the recommendations made (Finn & Hughes, 2008; Cossar et al, 2013; Donaldson & Irwin, 2017). A more theoretically informed evaluation of the client–counsellor relationship draws on psychodynamic traditions and involves evaluation of the therapeutic alliance (also referred to as the therapeutic bond or working alliance) – a negotiated, collaborative relationship entered into with both sides hoping to create change. A couple of studies demonstrated that it is possible to establish a therapeutic alliance in telephone and online counselling (King et al, 2006; Hanley, 2009). However, there are also criticisms of using service satisfaction as an outcome, as on its own satisfaction does not necessarily mean any tangible change to the circumstances (Mishara et al, 2007).

While satisfaction with the service was considered as an outcome on its own, some studies also sought to establish its effects on other helpline outcomes. For example, an evaluation of the Dutch Kindertelefoon, where support is delivered via telephone and online chat application, explored how the change in wellbeing and perceived burden of the problem correlated to user's satisfaction with the conversation with the helpline professional (Fukkink & Hermanns, 2009). Based on scores before and after the conversation, the study found that an increase of wellbeing and a decrease of the burden of the problem was associated with greater satisfaction with the service. Hence, user satisfaction can also function as a pathway to achieving better outcomes in other areas.

### Case study: Child Helpline International in Denmark (BørneTele-fonen), the Netherlands (De Kindertelefoon), Belgium (Awel) and Italy (Ill Telefono Azzurro)

A comparative study of the four child helplines that are part of the Child Helpline International network used three different collection points for measuring children's satisfaction with the service provided (van Dolen & Sindahl, 2017b). It employed questionnaires filled out by the counsellor and by the child immediately after the counselling session, and a follow-up questionnaire sent to the child two weeks after the counselling session (via email or SMS). The study found that immediately after the service, 80 per cent of children felt that they have been heard, 77 per cent were satisfied with the helpline, 66 per cent found the session helpful and 74 per cent would recommend the helpline to a peer. The children who used the helpline's chat service had the highest scores and those using the phone the lowest. The perceived helpfulness of the session was tested two weeks later and over half of the children (53 per cent) still thought that the session was helpful.

### 4.1.3 Changes in feelings and attitudes

Providing emotional support is one of the key aspects of the work of helpline staff, and it is unsurprising that most of the studies we reviewed offer some evidence of changes in feelings and attitudes. Arguably, one of the most well-established areas of outcome evidence is changes in children's feelings and attitudes. This relates both to the presence of a substantial amount of research to demonstrate effectiveness as well as the lack of criticisms of this approach – in contrast to the ones discussed so far. Overall, the findings demonstrate that, regardless of the channel, helplines have a positive effect on children's feelings and attitudes. This might involve increased positive feelings like hope, confidence and self-efficacy, reduced experiences of negative emotions like distress, isolation and perceived burden, or both. The importance of addressing feelings and attitudes is also demonstrated by evidence that in some cases help-seekers put more emphasis on the emotional processing than problem-solving (Law et al, 2015), or that the most positive effect from the service is seen in relation to feeling listened to and understood (Coveney et al, 2012).

Most of the studies measure a range of emotion-related outcomes. For example, an evaluation of Kids Help Phone (Canada) uses a pre- and post-test design with an online questionnaire to evaluate level of distress, perceived difficulty of the problem and self-efficacy using client self-report scales (Law et al, 2015; Haner & Pepler, 2017).

Changes in level of distress are measured by how upset the user feels about their problem, how strongly they are feeling their emotions at that point and how stressed out they feel. In a similar pre- and post-service design evaluation of the Australian Kids Helpline, King and colleagues (2006) found that the telephone counselling had a much stronger positive effect on distress reduction than online counselling, which they explain with the more intense communication of the telephone interaction. This corresponds to the findings about Kids Help Phone (Canada) where the highest effect was observed in relation to reduced distress levels of the telephone helpline, while the chat service was less efficient in that aspect.

### Case study: De Kindertelefoon, the Netherlands

Fukkink and Hermanns (2009) conducted a comparative study of the telephone service and the confidential one-on-one online chat of the Dutch child helpline, De Kindertelefoon. One of the most comprehensive and high-quality evaluations in the sample, the research comprised a concise pre-test, a post-test and a follow-up test (via an online questionnaire) one month later. The study measured wellbeing (Cantrill ladder changed to a 9-point scale to be used via a phone), the perceived burden of the problem and the degree of socio-emotional problems (via a Strengths and Difficulties Questionnaire [SDQ]) using standardised measures. The findings demonstrate that children who used both the telephone helpline and the online chat reported an increase in wellbeing and a decrease in the perceived burden of the problem after the session, with online users reporting better outcomes. The improvement was largely maintained at follow-up one month later. These positive outcomes, however, did not apply to each individual child. During the follow-up, one in four children who had reported a severe problem when they contacted the helpline (defined as a score of 6 or higher) thought that the problem was equally bad or worse. Another shortcoming is that the study had a low overall response rate (66,981 children contacted the service and only 12,873 answered the pre-service questions) and high attrition rate (only 223 children took part in all three data collections).

### 4.1.4 Changes in awareness and knowledge

An important aspect of the work of helplines is to enable users to understand better their situation and the possible solutions of the difficulties they are experiencing, including where to find further help, if needed. Some of the evaluations in the sample try to capture these changes in awareness and knowledge by looking at how valuable the provided information is for solving the problem, whether the child has a clearer understanding of their situation, and if they are more aware of the available resources. For example, van Dolen and Sindahl (2017b) assessed the child's instrumental empowerment resulting from the service evaluated at the end of the session by both the child and the counsellor. They found that after talking to the helpline, over half (54 per cent) of the children had an idea of what to do and this effect was maintained two weeks later. Children who used the chat and email services had higher rates, and those who used the telephone helpline had the lowest. Over two-thirds of the children (67 per cent) also found that they had been given plenty of information and advice, with those accessing the service having the highest score, followed by the chat users and the telephone helpline showing the lowest scores.

Other studies in the sample include evaluation of changes in relation to knowledge of resources and support available or knowledge of how the system works. The Prison Advice and Care Trust Helpline, for example, specialises in supporting families and friends of prisoners including providing information about visits and contacts with the prisoner or the institution, non-statutory and statutory sources of support, and the terminology and processes used in the criminal justice system (Sharratt, Porter and Truman, 2014; Goldsmith, 2017). The independent evaluation of the service including a user survey and interviews demonstrated that the helpline offers useful information to the users – 80 per cent of respondents strongly agree or agree that staff provide the information or advice needed to help them solve the problem. Finally, scaling up to the social level, changes in awareness and knowledge may refer to public attitudes towards children's issues (support, risks, harm) and awareness of the available services, such as helplines.

### Case study: Kids Help Phone, Canada

Law and colleagues (2015) compared the outcomes of the telephone and chat helpline services provided by Kids Help Phone, Canada. Using a pre- and post-test research design with questionnaires filled in by 232 telephone users and 129 chat users, they explored the user expectations as well as outcomes like increased problem clarity, increased self-efficacy, reduced distress, perceived problem difficulty and increased hope. The study found that, overall, 69 per cent of callers or chatters saw their problem more clearly after speaking with a counsellor, with higher rates of improvement observed for chat support. Still, at the end of the session the average rating for problem clarity showed statistically significant improvement for phone support. Children who contacted the helpline via the chat had an overall lower score for problem clarity before using the service but showed a better improvement after the service.

*For further details on this case study, see also Wilson and Haner (2012).*

### 4.1.5 Skills development

There is an overall agreement in the reviewed studies that the process of helping should involve supporting the child to articulate a strategy of solving the problems that they are facing or even to formulate clearly what changes are necessary to improve their situation (Hepburn, 2005; Danby & Emmison, 2012; Potter & Hepburn, 2012). While this is seen as an important outcome for the child, skills development is not always clearly identified as a measurable outcome, possibly due to the difficulty of creating such changes in a short intervention and measuring them. A possible solution of this dilemma is offered by Mishara and colleagues (2007) in their study of the US suicide prevention helpline, Hopeline. The study involved analysis of the helping process and reactions of the caller by listening and coding 2,611 calls to the helpline. Mishara and colleagues (2007) analysed if the process of helping followed established practices that, they argue, have been demonstrated to lead to positive outcomes for the user, such as collaborative problem-solving, asking factual questions about the problem and precipitating events, suggesting a plan of action, and offering referrals. The same research design was also used by Mokkenstorm and colleagues (2017) in the study of the Dutch crisis chat service 113Online, where the visitors were rated in relation to helpless/resourceful, being confused/decided, etc. Another example of a possible way of exploring skills development is offered by Morgan and colleagues (2012) who did an evaluation

of mental health helplines and discovered that the majority (69 per cent) of callers felt that contacting the helpline enabled them (a lot or somewhat) to cope better with their condition, to stay more focused and be more in control. However, these examples also demonstrate that skill development is certainly hard to establish in a helpline environment, and studies offering robust evidence-gathering models for demonstrating effectiveness in relation to skill development are not found in the evidence review.

### 4.1.6 Changes in individual behaviour and practice

Demonstrating changes in individual behaviour and practice seems to be even more challenging to evaluate, and fewer studies tried to do this. Methodologically, this outcome requires some follow-up evaluation as it takes time for changes to occur. A possible proxy is measuring intentions to make changes, for example, the likelihood of seeking support from other services as a result of the information from the helpline (see, for example, Goldsmith, 2017). This, however, does not offer strong support for outcomes related to changes in behaviour, as it does not capture actual practices. Another possibility to explore changes in behaviour and practice is examining returning callers but there was not enough evidence from the reviewed studies to evaluate this approach. The evaluations were mostly based on the last time a child contacted the helpline, with some studies including 'first contact from this IP address' among their inclusion criteria (see Mokkenstorm et al, 2017). The few studies looking at repeated clients often acknowledged that this might be seen as a sign of helpfulness and greater benefit or the opposite – that the issue was not resolved effectively (Mishara et al, 2007; Wilson & Haner, 2012).

The changes of individual behaviour and practice are captured better by evaluations of services that have both a helpline and a face-to-face element. For example, Shelter Children's Services incorporates a helpline, a web service and a national network of local centres. The evaluation of the service was able to trace changes in individual behaviour related to educational outcomes (improved attendance, ability to engage with the educational process, school attainment, reduced bullying), emotional and social wellbeing outcomes (improved emotional literacy, increased confidence, improved behaviour, improved parental resilience and family functioning, improved social skills and increased participation) (Shelter, 2011). However, there is little evidence on how a helpline service can create change to behaviour and practice on its own.

### Case study: YoungMinds, UK

Donaldson and Irwin (2017) used an online survey and a follow-up telephone interview with parents who had used the YoungMinds helpline. The survey was completed by 273 service users, recruited by email: 57 per cent of them had used the phone frontline service, 43 per cent the email frontline service and 33 per cent had received a telephone call-back. Following the survey, 19 parents undertook a more in-depth telephone interview. The research focused on the advice and information parents received and if they acted on the help; what changes to individual and family behaviour followed; the outcomes of the telephone service compared with the email frontline; and if there was any additional impact of the call-back service. The study offers an interesting approach to evaluating the changes in individual behaviour as it maps the outputs (advice given by the helpline) to outcomes (taking action as a result of the advice) and effectiveness (the advice received was helpful for those who followed it). The respondents of the study also evaluated the extent to which their behaviour changed and how much of this change was a result of the service. Changes were reported in relation to the more effective use of mental health services (30 per cent report change), improved relationships with family (35 per cent), better child behaviour (36 per cent), and better relationships with the child (40 per cent), with 89 per cent of parents attributing at least some of the change to the helpline.

It is unclear if the measures can be applied to children, particularly given the complex attribution of effects to the helpline. The design also requires following up the service users, which would not be possible in cases of anonymous users. Still, the model offers an interesting approach to evaluation of longer-term behaviour change.

### 4.1.7 Changes in family practice and social behaviour

The further we move away from the individual level, the less evidence exists in relation to measurable outcomes and a small number of evaluations that explore these outcomes. Changes in family practice and social behaviour is an area that is lacking in the evaluation research. This is particularly surprising in relation to peer relationships, where we were expecting to find more evidence. The gap might be related to difficulties in measurement. The existing evidence related to family and behavioural outcomes, we found, usually refers to interventions that have a face-to-face element or that are targeted at children as part of a family (Shelter, 2011; Starks et al, 2012; Daly et al, 2015). The case study above, on Donaldson and Irwin's (2017) evaluation of YoungMinds in the UK, is one of the examples where

an intervention aimed at parents explores how the helpline has changed practices within the family. The lack of similar examples in relation to child helplines demonstrates the need to make more efforts to ensure that children are perceived as active players in their family and social surrounding.

### 4.1.8 Changes in organisational practices and capabilities

Another possible area of outcome research addresses changes in the helpline itself, for example in its organisational practices or capabilities. There are a small number of studies in the sample that focus on this, evaluating procedures, competence, evidence-gathering techniques and collaboration with stakeholders (Kitchingman et al, 2015; Dinh et al, 2016; Child Helpline International and UNICEF, 2017).

One example in this area is the evaluation of the Australian crisis support line that measured the counsellors' readiness for competent service delivery. Kitchingman and colleagues (2015) used vignettes to introduce the stories of three types of callers – suffering from severe suicidal ideation, a major depressive episode, and acute general anxiety. Using the Telephone Crisis Support Skills Scale (TCSSS) and self-reported intentions, the research identified the counsellors' skills for working with different types of callers. Such methods can be used to demonstrate effective training and quality control of the work of the helpline.

Two other studies, of the helpline networks of Child Helpline International and Insafe, explored organisational procedures, capabilities and effectiveness in responding to online risks for children (Dinh et al, 2016; Child Helpline International and UNICEF, 2017). For example, Child Helpline International administered a questionnaire to 17 helplines in their network aiming to determine current capabilities in relation to online sexual exploitation and abuse (Child Helpline International and UNICEF, 2017). The study looked at the priority of online protection in the helpline operations, strategy and planning; knowledge of the respective legislation and involvement in providing inputs for improvement of the legal framework; cooperation with referral networks; presence of protocols for handling and referring cases of online sexual exploitation; staff training on online protection issues; and efforts for raising social awareness of online protection. The research identified strengths and weaknesses in the helplines' ability to respond to online child sexual abuse and exploitation and showcased best practice examples to guide future performance. Similarly, the Insafe study used the findings to make recommendations about the support required and the resources needed to respond effectively to the new and emerging risks that internet use may pose for children.

Finally, there are individual cases of studies exploring the change of organisational procedures of other organisations – for example, a peer support emailing service enabling a local school to tackle bullying (Hutson & Cowie, 2007). Overall, changes in organisational practices and capabilities are addressed but they are relatively marginal in the sample of studies included in the review.

### 4.1.9 Changes in cross-organisational practices and community dynamics

While the importance of collaboration and coordination of efforts between agencies and stakeholders within the sector is recognised in the reviewed studies (Hutchinson, 2012; Nicholas & Broadbent, 2015; Dinh et al, 2016; Child Helpline International and UNICEF, 2017), outcomes at the cross-organisational and community level are rarely discussed and the evidence provided is particularly limited. When such evidence is presented, it is mainly in the form of case studies where the discussion of effectiveness is not necessarily linked to the measurement of clearly defined outcomes and not subject to an evaluation. Still, there are some motivating examples that demonstrate the need to expand both the efforts to work in this field and to pursue more comprehensive and robust ways of demonstrating effectiveness. For example, an ethnographic study on patients' organisations demonstrates the importance of the work on building relationships across sectors, easing the flow of information, and enhancing "the collective intelligence of the entire…system" (Nicholas & Broadbent, 2015: p.4).

Another relevant study is the evaluation of the Safe@Last support service for children under 17 that comprises a helpline (telephone, web and text services) as well as a range of local services, such as a refuge, preventative education and street work (Starks et al, 2012). Alongside case studies of children and the service, the evaluation included a 'partnership case study' comprising interviews with key stakeholders at both a strategic and operational level, such as representatives from local authorities working with Safe@Last, representatives from children's services, the police, schools and social services. While the evaluation included robust child outcome and cost-effectiveness measures, changes in cross-organisational and community dynamics were not linked to particular outcomes. The effectiveness at this level is demonstrated by a narrative about the partnerships (such as schools, youth clubs, social care residential homes and community centres, neighbourhood police officers and drug prevention workers) and the number of children reached by the service – 959 via the detached street work and 33,000 via schools. The study also mentions the efforts to ensure consistent support from other key services, such as the police and social services, which also required a contribution to local policies and procedures, such as runaway

protocols and local safeguarding strategies. Based on the case study work, Starks and colleagues (2012) concluded that, to be effective, the support needs to be part of a broader community response and to work in partnership with other services, such as schools, social services, mental health and family support services.

### Case study: Netsafe New Zealand

During the expert consultation accompanying this evidence review, Neil Melhuish from Netsafe New Zealand discussed with our team the dilemma their helpline is facing. Children are underrepresented among the people who contact them for help, and Netsafe has been trying to answer the question, "Should we work to attract children to contact us directly, or to strengthen their support networks?" The way ahead seems to be doing both: "We can improve our offering to children, for example by piloting a messaging channel as suggested by children in the recent *Insights into digital harm* report by the Ministry of Women and Netsafe (2017). At the same time, we try not to lose sight of the need to raise the capability of the people and institutions that support children. Netsafe has focused on reaching out to the adults in children's lives – parents, teachers, community police, policy makers, other helplines – in an effort to improve the support networks that can mediate risk, increase children's help-seeking efficacy, and adults' preparedness to identify and respond to problems."

*Provided by Neil Melhuish, Netsafe New Zealand.*

While these examples demonstrate the importance of working at cross-organisational and community level and the numerous efforts of child services in this direction, they also highlight the lack of robust evidence and examples of good evaluation mechanisms with clearly identified outcomes and measures, which might be expected considering the defused nature of this type of outcome.

### 4.1.10 Changes in policy, protection and society

Similar to cross-organisational and community outcomes, changes in relation to policy, protection and society are rarely the main focus of the reviewed studies. The evidence is mostly sporadic and there are no good models that can systematically map the areas of influence and the relevant outcomes. A report by NSPCC (2016) reviewing Childline's 30 years of activity is an example of a good overview of possible key areas of influence, including campaigning and helping to shape legislation; raising awareness; providing evidence, resources and tools; working collaboratively; and having international influence.

The report also gives examples of specific outcomes or outputs under each of these areas, for example, successfully campaigning with other organisations for the promotion of the safeguarding and welfare of children in the Children Act 1989 (under legislation change); running a Fight Against Porn Zombies campaign to help build children's awareness of online pornography (under awareness); submitting key evidence into the National Inquiry into Self-harm in 2004 (under evidence); and co-authoring the first good practice guidelines for telephone helplines for children in Europe (under international influence). This approach demonstrates the effective use of case studies and examples to establish the contribution of Childline. This type of evidence, however, has its limitations, as it does not rely on clearly identified criteria for effectiveness or on an independent evaluation of effectiveness (e.g., in relation to the awareness campaign).

Some examples of more systematic evaluation of effectiveness at the social level can be found in a small number of studies involving cost-benefit analysis. These studies concern outcomes like reduced social costs associated with the long-term effects of child sexual abuse, anti-social behaviour, the reduced support needs of children in crisis, prevented accidents and emergency visits or paediatric ward stays (Shelter, 2011; Starks et al, 2012; Bowles, 2014; Calderón et al, 2017; Cullen et al, 2017). While these studies have robust evaluation designs, they are usually issue-specific and capture a small area of benefit at the social level.

Overall, the studies that explore the outcomes at the level of policy, protection and society are small in number and mostly have methodological limitations as they are issue-specific or rely only on case study evidence. This identifies a need for more comprehensive mapping of social level outcomes, which are measurable and correspond to the areas of work of child helplines.

The review demonstrates that helpline outcomes span from the individual to social level, but there is considerable disparity in the amount of existing evidence available. Child outcomes have been subject to more robust evaluations, particularly in relation to service satisfaction, changes in awareness and knowledge and changes in feelings and attitudes. Much less evidence is available on outcomes occurring at the interpersonal, inter- and intra-organisational, communal and social levels. In the next section, this report offers a more comprehensive overview of the different approaches to effectiveness and the types of evaluation designs.

## Case study: Insafe helplines, Europe

Dinh and colleagues (2016) carried out an evaluation of the effectiveness of 31 Internet Safety (Insafe) national awareness helplines that aimed to explore the main services that helplines provide; how helplines measure impact and effectiveness; the key emerging issues and challenges that helplines are dealing with; and the strategies that helplines employ to deal effectively with these issues. The research used mixed methods comprising a detailed comparative study of four helplines – the Safer Internet Centre (Belgium), BeSmartOnline! (Malta), Childline (Ireland) and Sigur.info (Romania) – including interviews and focus group discussions with key helpline personnel and national stakeholders, as well as looking at output data and analysis of documentation (such as annual reports). In addition, a survey of 24 Insafe helplines was conducted.

In terms of outcomes at the social level, the study used mostly interviews to explore staff perception of how well known the helpline is, to gather information about engagement of the helpline in various public awareness activities and if any measures of effectiveness are used, as well as perceptions of which activities are most effective. The participants were also asked to share their opinion on the level of awareness and cooperation between the helpline and industry, government and other agencies. The interviews with staff also elaborate on the public awareness strategies used (such as door-to-door visits, distribution of flyers, collaboration with advertising agencies, media participation, support from government and public figures). In addition, the survey provided information on the proportion of helplines that use public awareness activities (advertising on national media, online advertising, advertising on own website, information sessions for the community, at schools, at public events, etc.).

However, measures of effectiveness were not discussed in the report. The only exception is an awareness campaign by the Romanian helpline Sigur.info that used the following indicators: increased number of website visits and growing social media popularity (Facebook 'likes'), and direct reach of children, parents and teachers by the services.

# 5. Evaluation of effectiveness

## 5.1 Approaches to effectiveness

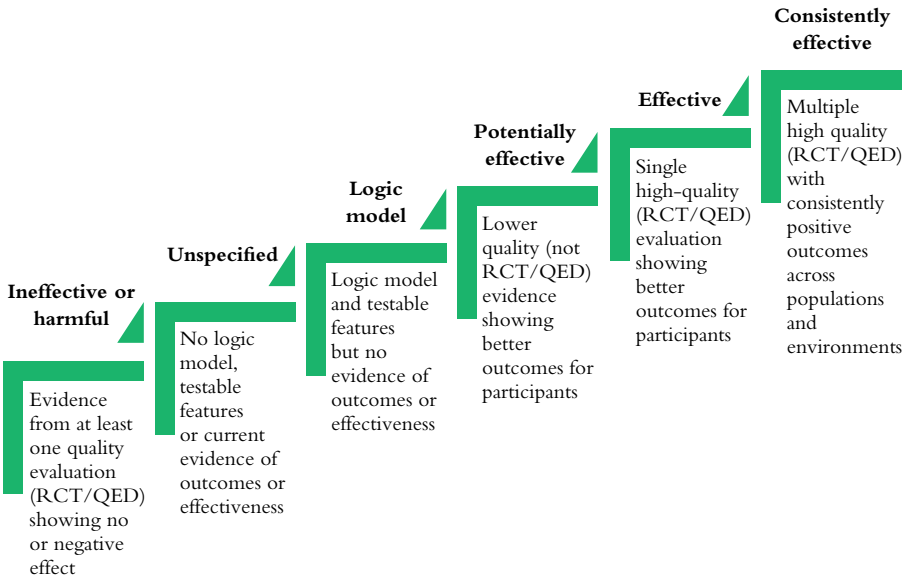
The systematic review demonstrates that the helpline sector faces substantial challenges in evaluating its effectiveness. This is related both to the struggles of individual helplines to carry out high-quality evaluations of their services, as well as to the lack of established and recognised models and outcome measures within the sector. Effectiveness is, therefore, used to refer to various types of evidence related to the work of the helpline, and the different approaches can be grouped under the following:

- ***Effectiveness model***: only a few studies in our sample refer to a model of effectiveness that was used to guide the evidence gathering (see Shelter, 2011; Daly et al, 2015; Law et al, 2015; Child Helpline International and Oak Foundation, 2017). The commonalities in the models we found relate to the mapping of inputs, outputs, planned and achieved outcomes and some longer-term benefit from the service (or 'impact'). Effectiveness models might involve some underlying Theory of Change (Rogers, 2008; Vogel, 2012; Morton, 2015) or at least some mapping of the pathway of the desired change and the stakeholders involved.
- ***Measuring outcomes***: a substantial number of studies identify service outcomes and use empirical evidence to demonstrate that these outcomes have been achieved. These studies usually involve some form of formal evaluation and have a relatively high reliability, if robust research designs have been used.
- ***Studies of the process of intervention***: some studies focus on demonstrating good practice during the intervention process drawing on theoretical models or prior empirical evidence of what characteristics of the service can lead to better outcomes. While these studies do not measure the effects on the user, they draw on solid evidence of what constitutes good practice within the sector.
- ***Evidence on outputs and activities***: another approach to effectiveness focuses on presenting evidence about the activities and outputs of the helpline, capturing the services provided and beneficiaries affected. This approach usually involves some demonstration of the need of the service provided, for example, via the use of case studies. The strength of this approach is the relatively easy evidence gathering (usually part of the working process of the helpline) and ability to provide regular updates.

- ***Combined approach***: it is not uncommon for studies to draw on more than one of these approaches to capture the complexity of the helpline work and to demonstrate effectiveness more efficiently.

While these approaches are observable in the sample studies included in the review, there are very few overviews that use a framework for intervention effectiveness combined with matching types of evidence of effectiveness. One of the exceptions is the work of Cullen and colleagues (2017), who offer a useful classification of effectiveness in their adaptation of the model originally developed by the Early Intervention Foundation (see Figure 2 below). Their model suggests a five-step hierarchy of effectiveness starting from 'ineffective or harmful' and leading to 'consistently effective' interventions that are able to demonstrate high-quality evidence for reliably positive outcomes across different groups and contexts.

## Figure 2: Intervention effectiveness and outcome evidence




Notes: RCT – randomised controlled trial; QED – quasi-experimental design.
Source: Based on Cullen et al (2017)

Applying this model, we would suggest that half of the helpline evaluations in our sample are able to demonstrate effectiveness of the service. Of the 51 studies using primary research, 24 have a quasi-experimental design and one has an experimental design – hence, presumably, consistent effectiveness. The other half of the primary research studies can only demonstrate potential effectiveness of the services evaluated based on the methodological approach (exploratory, case study, observational or ethnographic, narrative or discursive and interview/focus group) (see Table 1).

*Rather than suggesting that the helplines are ineffective or that the used methods of evaluation are poor, this finding should be used as a prompt to question the extent to which established models of intervention effectiveness and reliability of data can be applied to the evaluation of helplines, and particularly to those who work with children. Defining the scope of what is methodologically and ethically achievable can help to set realistic and feasible evaluation targets within the helpline sector (Mental Helplines Partnership, 2003).*

## 5.2 Evaluation designs

The systematic evidence mapping discovered a number of evaluation designs that have been used by helplines to gather evidence of effectiveness. To systematise the findings, we use the model offered by Child Helpline International and Oak Foundation (2017), which presents one of the most systematic classifications of the different evaluation designs for demonstrating helpline effectiveness that the review came across. We slightly revised the original model in light of the review findings. The classification includes the following evaluation designs: output statistics; service evaluation; surveys; web feedback and evaluation; focus groups, panels and interviews; self-assessment of call by counsellor; case reviews; referral feedback; 360-degree evaluations with stakeholders; community opinions, awareness, attitudes and help-seeking sampling; and cost-benefit analysis.

### 5.2.1 Output statistics

A key source of evidence for the work of the helplines is output statistics that are used in almost all of the studies included in the review. Output statistics are often presented with some element of comparison – to previous years, between the different branches, between similar helpline services in different countries, or between different channels of the same service. Most helpline studies provide a comprehensive overview of the calls including number of received and answered calls; average duration of calls; average number of calls per caller; proportion of new callers; and follow-up services provided (where this was an option). The services also hold data on the key issues raised by the users, which vary based on the service provided, but include areas like self-harm, health, education, eating problems, suicide, problems with friends or partner relationships, bereavement, gender or sexual identity (Childline, 2012; 2016). Issues related to anonymity of the callers sometimes pose challenges to the collection of user data, but most helplines seem to record some basic characteristics of their callers (such as age, gender, ethnicity and area of residence). In some cases, output data is presented in comparison to target populations, and particularly in relation to vulnerable or seldom heard groups or in relation to data from other helplines within a network

(Gallagher, 2013; 2014). In some cases, the output data includes indicators like capacity used, which compare the time used by helpline staff to the time for which the lines are estimated to have been open (Bowles, 2014). For services with digital platforms, the output data also includes site visits, online counselling sessions and use of particular parts of the platforms.

### 5.2.2 Service evaluation

Service evaluation usually includes inviting the participant at the end of the session to describe their experience or satisfaction with the service. The sample of evaluations included in the review demonstrate various ways in which this type of evaluation can be conducted – automated phone or online evaluation, the counsellor conducting the evaluation at the end of the session or the user being transferred to a special evaluator. Service evaluation is usually used to capture the short-term effects of the service and utilises the benefits of discussing the support intervention immediately after it happens, reducing confusion about what is being evaluated and memory gaps.

There are some shortcomings, however. In addition to being sensitive to collect, the data gathered via this method is susceptible to bias at two levels. First, participation depends on opting in to do the evaluation, which makes the data vulnerable to selection bias. Not having information about the people who decide not to participate makes any judgement about similarities between the participants and overall group of users hard to make. Second, the measurement of outcomes is based on user self-evaluation and might be subjective – for example, the users' initial expectations of the service would influence the way they evaluate their outcomes. As the evaluation only captures the post-service state and there is no baseline data, establishing the reliability of the self-evaluation is difficult. Some of the studies in the sample attempt to balance this by asking the users about the extent to which the service met their expectations and exploring how much this influenced the outcomes and plans to follow the advice given. For example, in evaluating how satisfied the children are with the Kids Help Phone (Canada), Law and colleagues (2015) included an open-ended question in the survey asking "What did you want to get out of your call with a counsellor today?" and the responses were coded into three categories: emotional processing/management or talking; issue-based discussion or problem-solving; or other. They also recorded whether the child was likely to use the service again and recommend it to others.

Another methodological approach of service evaluation that we came across during the review is a silent monitoring study (Mishara et al, 2007). This type of study involves a trained researcher listening to helpline calls, with prior agreement from the users, and noting information about the nature of the help provided and callers' reactions using standardised instruments. The assessment is taken twice – at the beginning of the call and towards the end – and measures the short-term effects on the caller using a 26-item scale with 16 items to indicate successful outcomes (empathetic understanding; respect of the caller; collaborative problem-solving) and 10 that indicate non-successful outcomes (being directive; judgemental; not well informed, etc.), rated on a scale from 1 (strongly disagree) to 7 (strongly agree).

### 5.2.3 Surveys

Surveys can offer a good way of obtaining information for evaluation purposes, especially if they are administered on a regular basis, for example, at least once a year (Dinh et al, 2016). The use of questionnaires is fairly common in the sample of evaluations included in the review and is most often employed in quasi-experimental or exploratory designs. While experimental designs offer the highest level of reliability of effectiveness evidence, we only found one randomised controlled trial study and it was of an adult hotline. When used with service users, the surveys often evaluate a number of outcomes, such as service satisfaction, changes in user's behaviour, knowledge or emotional state, as well as to establish the profile of the user. The ideal method of using a survey to establish helpline effectiveness is to employ a pre- and post-service questionnaire that can capture the changes and measure the outcomes more reliably. For example, two different online services for children – Kids Help Phone (Canada), which offers online chat counselling, and Kooth.com (UK), an online therapy service for children – were both evaluated using pre- and post-questionnaires to measure outcomes, such as mental health and the client–counsellor relationship. In some of the reviewed studies, only a post-service questionnaire was used, which invites the participants to evaluate the outcomes of the service. For example, an independently conducted evaluation of the US helpline iCanHelpline.org consists of a customer satisfaction survey administrated via the phone or online, and focuses on satisfaction with the service, key issues raised, technical assistance provided (e.g., deleting inappropriate online content) and likelihood of recommending the service. There are also examples of using questionnaires with helpline staff to establish competence and gaps within the helpline itself (Dinh et al, 2016; Child Helpline International and UNICEF, 2017).

## Case study: Children's Rights in Society (Barnens rätt i samhälletm, BRiS), Sweden

A study of the Swedish helpline Barnens rätt i samhället (BRIS, Children's Rights in Society) by Andersson and Osvaldsson (2011) used pre- and post-service questionnaires to establish the effectiveness of the online services (chat, online forum and email). The repeated measures included service satisfaction (being listened to, being taken seriously, the service is helpful) and improved wellbeing (on a scale from 1 to 9, where 1 represents 'very poor' and 9 'great'). Other measures included increased problem clarity (reporting they know what to do with the problem as a result of the service; understanding and being better able to explain what the problem is), decreased problem difficulty and increased self-efficacy (feeling more self-confident). The effectiveness was demonstrated by the positive outcomes immediately after the study and 10 days later. The study found that the strongest effect immediately after the service use was on wellbeing and knowing what to do about the problem. However, the most long-lasting effect with the smallest decrease over the 10 days after the service was perception of problem difficulty, understanding what the problem is and being able to explain what the problem is.

## Case study: Suicide Hotline, US

An example of an experimental study (although not with children) is the evaluation of the effectiveness of the US Suicide Hotline that involved a randomised controlled trial design (Rhee et al, 2005). All callers to the crisis hotline are assessed by a crisis worker for suicide risk and only callers who are not currently in therapy, who are at no or low risk for suicide, who have no indications for psychiatric referral and who express interest in beginning phone therapy were included in the study. Those who agreed to participate were randomly assigned to one of three treatment conditions: Solution-Focused Brief Therapy (SFBT), Common Factors Therapy (CFT) or a Waitlist Control (WC). All participants completed pre- and post-treatment questionnaires including standardised measures of depression and life satisfaction, as well as a pre- and post-therapy interview with an independent therapist. A range of outcomes was assessed including hostility, anxiety, depression, interpersonal sensitivity and life satisfaction. Effectiveness was measured based on the scores from the self-assessment and the therapist rating. The study found that for 10 of the 14 outcome measures, clients receiving therapy reported better outcomes than clients placed on the waiting list. While we could not find a study that used this design with children, over 15 per cent of the participants in this study were young adults (under the age of 25), which demonstrates potential applicability to child helplines.

### 5.2.4 Web feedback and evaluation

With the increase in digitally mediated service provision for children, it is hardly surprising that new research methods to capture effectiveness have come to the fore. Using websites to seek feedback and evaluate the service is one such example. Web feedback and evaluation took several forms in the studies in the sample – filling in an online form with feedback about the service, analysis of online forum content and hosting an online survey. While the studies in the review usually report site visits and use of online content, some also discuss using websites for receiving feedback about the service. In her evaluation of the US YooMagazine health website for children, Rogers (2008) used web feedback as one of the methods. After the third visit on the website, the users were asked to provide feedback on the usefulness of the information on the website. They also identified if they would recommend the website to a friend, if the website was helpful personally to them and if they had learned something new from it.

There are also examples where evidence of effectiveness is sought by analysing the web content itself. For example, in the study of a range of child abuse and neglect services in the UK, Cossar and colleagues (2013) did a content analysis of an online peer support site where children post and respond to problems involving abuse and neglect. By exploring how children understand and conceptualise abuse and neglect and the barriers surrounding recognition and telling, the study offers insights into the ways in which the online forum is enabling young people to receive support from each other. The findings suggest that children receive emotional support and reassurance from their peers that they are not alone and are being encouraged to seek formal help.

In addition, the digital environment has offered new opportunities to seek follow-up feedback about the service in attempts to capture longer-term effects or to conduct evaluations via an online questionnaire in cases where anonymity and confidentiality do not allow the direct recruitment of participants. For example, Coveney and colleagues (2012) hosted an online questionnaire on the project website for their study evaluating the effectiveness of the Samaritans' telephone and email support service, which allowed the recruitment of a wide range of different callers. The sample was collected over a year and comprised 1,309 analysable cases, of which 10 per cent were of children under 16, and 36 per cent were between 16 and 24 years of age. To reach callers who were less familiar with the internet, the survey was advertised through local and national media and community networks and a paper copy was available on request. The online survey was able to demonstrate that the respondents felt more positive than negative after their last contact with the Samaritans, with highest median scores for feeling listened to and feeling understood.

Because of its design, the study was also able to demonstrate how the Samaritans compares with other sources of support. The constraints of this approach concern the inability to employ a sampling strategy or to recruit a representative sample that exposes the results to self-selection effects.

### 5.2.5 Focus groups, panels and interviews

A range of qualitative methods is used in the reviewed studies as a way of capturing the richness of the lived experience of seeking helpline support. Focus groups or interviews were conducted with children, staff, family or community members and various stakeholders, both independently and as part of mixed-method research designs. They can be particularly useful when trying to establish hard-to-quantify effects, such as preventative work. For example, Calderón and colleagues (2017) conducted interviews with child service users, their parents and service providers in an attempt to estimate the preventative effect of the helpline and the face-to-face service for out-of-hours support for children with mental health problems. The interviews demonstrate that the service is able to offer early intervention and support to children and their families to better manage the emerging crisis in a home environment, preventing the need for other services. Focus groups or interviews can also be used to establish the operational dynamics and existing protocols of the helpline and to identify difficulties and good practice solutions (Statham & Carlisle, 2004). Another reason to use this type of qualitative method is that it allows children to voice their concerns and provides an opportunity for their stories to be heard (University of Central Lancashire and NCB Research Centre, 2016).

### 5.2.6 Self-assessment of call by counsellor

Self-assessment by the counsellor is used in a number of studies in the sample, and there is an overall agreement that the counsellors are in a good position to evaluate the service and the outcomes. Different approaches are employed including rating the helpfulness of the session and evaluation of user outcomes (Finn & Hughes, 2008). Karver and colleagues (2010), on the other hand, conducted a study that aimed to examine whether the assessments of the helpline counsellors were reliable. The team analysed the judgement of risk for suicide-related behaviour by 35 US and Australian-based counsellors who all looked at 45 child cases over a six-month period, exploring whether the counsellors agreed and whether their judgements were consistent with the children's actual behaviour in a six-month period following intake. Karver and colleagues (2010) found that helpline counsellors generally agreed and were accurate in their judgements – there was 80 per cent correct classification in identifying children who engaged in suicide-related behaviour during the six-month follow-up period.

## Case study: National Sexual Assault Online Hotline (NSAOH), US

The evaluation of the National Sexual Assault Online Hotline (NSAOH), an online hotline for rape and abuse, involved collecting feedback from users, volunteers and supervisors (Finn & Hughes, 2008). The online service is run by volunteers, and while the chat sessions are based on the visitors' needs, they usually include support, information and referral, and problem-solving. If a volunteer is having difficulty in a session, the system allows them to get help from an online supervisor through a chat. The supervisor can send messages directly to the volunteer and to the visitor. Supervisors may also decide to intervene if necessary, while routinely monitoring volunteers' sessions as part of quality control.

The self-assessments of the volunteers were gathered via an online form at the end of their shift about their last session. This method was used in order to reduce volunteer time in answering surveys and as it provided some randomisation of the data collection. The survey focused on the perceived helpfulness of the session from the volunteer's viewpoint, type of issues discussed, difficulties experienced with support services and types of service-related difficulties encountered in the session. Volunteers were asked to rate their perception of the helpfulness of the session on a 5-point scale from 1 (not at all) to 5 (a great deal). Approximately half of the volunteers (54 per cent) rated the session as very helpful (4 or 5), 25 per cent rated the session as somewhat helpful (3), and 21 per cent rated the session as not helpful (1 or 2). The mean helpfulness rating was 3.5 (SD = 1.1). Those who rated the session as less helpful had usually experienced difficulties during the session related to technical faults or the users stopping to communicate. The supervisor data was gathered at the end of sessions when they were called in by the volunteer (recording the type of difficulty encountered and the extent to which the problem was resolved successfully), as well as at the end of routine checks when volunteer chat sessions were randomly observed. Overall, supervisors believed that their intervention was very helpful for the volunteers and the users, with 75 per cent rating the intervention very helpful (4 or 5) to the volunteer, and 70 per cent rating the intervention as very helpful (4 or 5) to the users.

While the study could not match user evaluation with volunteer evaluations due to anonymity, this model offers the possibility of comparing two types of helpline staff evaluations, reducing the subjective bias.

### 5.2.7 Case reviews

Case reviews can be used in a number of ways, including for periodical internal assessments and staff training that can ensure quality control of the service (Child Helpline International and Oak Foundation, 2017). The sample of evaluations included in the review, however, comprise a few studies that can be seen as case reviews in a broader sense – they involve in-depth analysis of parts of the conversations between the counsellor and the child and demonstrate in various ways the effectiveness of the communication. Danby and Emmison (2012), for example, did a narrative analysis of parts of a telephone call to Kids Helpline Australia selected from a sample of 50 calls already audio-recorded for quality assurance and chosen randomly over a six-month period. In their in-depth analysis of the interaction, they illustrate the process of helping related to coming to a resolution to the problem in a way that is led by the child. In a study with a similar research design, Hepburn (2005) explores the process of providing evidence by the caller and the corresponding referral decision-making by the service provider, while Potter and Hepburn look at the response to crying (2005) and shared experiences of abuse (2012) in helpline support. These examples demonstrate in various ways the workings of the helping process and the principles of offering support and draw on established models of good practice within the sector, thus making claims about the effectiveness of the service.

Using case studies can also be seen as a type of case review. The case studies we came across are typically good practice examples and range from a small-scale case study of an individual service user to large-scale examples of the work of a whole helpline (Daly et al, 2015; Dinh et al, 2016; Child Helpline International and Oak Foundation, 2017). The studies that are better able to demonstrate effectiveness usually include some rationale for the selection of the case study and an explanation of the criteria for effectiveness (why this is a good practice example).

### 5.2.8 Referral feedback

Referral feedback involves getting in touch with the relevant agency a short time after referral and being asked to fill in an evaluation form indicating the follow-up treatment and any follow-up information that is known to them about the client (Child Helpline International and Oak Foundation, 2017). While most studies in the review reported on number of referrals, not many discussed any means of following up with the referrals. Among the small number of exceptions is the evaluation of Shelter Children's Services. A reciprocal referral system between Shelter staff and other children's services providers exists that allows for following up on the outcomes of the referrals and joint work on cases until their resolution (Shelter, 2011). However, for short interventions and signposting, Shelter does not record

the outcomes of referrals. Another approach to referral feedback is getting in touch with the service user and asking for evaluation of the referral helpfulness (Donaldson & Irwin, 2017) but again, this seems to be a rare practice in the sample analysed. Therefore, while the number of referrals is often reported, there are important evidence gaps in establishing the longer-term outcomes from the referrals mainly due to the methodological challenges of gathering referral feedback and the ethical challenges of working with vulnerable children. Nevertheless, there is evidence that child helplines make efforts to follow up, especially when there has been an intervention (as opposed to a referral) by the police, hospitals and child protection services (University of Edinburgh, Zimbabwe C and Zimbabwe UNCsFU, 2016).

### 5.2.9 360-degree evaluations with stakeholders

This approach allows a comprehensive evaluation of the helpline from the perspective of numerous stakeholders. It might involve a mixture of research methods, such as interviews, workshops, questionnaires with staff, children, partner organisations and management, site visits and public surveys to assess the overall functioning and effectiveness of the child helpline. Such comprehensive evaluations are expensive and there are not many examples of studies in the sample that included all of the above-mentioned elements, but there are some that comprised a number of them. An example of a comprehensive/360-degree evaluation is the study of Morgan and colleagues (2012) evaluating the effectiveness of mental health helplines and how they are perceived by stakeholders, including helpline staff and volunteers, health professionals and the people who call the helplines. The research had a mixed-method approach and involved: (1) staff interviews with helpline workers and managerial staff of nine mental health helplines; (2) interviews conducted with members of staff of community mental health teams (CMHT) (mental health nurses, community psychiatric nurses, team managers, intake workers, social workers, CMHT administrators and consultant psychiatrists); (3) a survey (administered online, via the phone or by post) with helpline callers; and (4) a survey sent out to general practitioners (GPs) via the National Opinion Poll Service to find out their views of helplines.

Even though there are not many examples of such comprehensive evaluations in the sample, this approach can be particularly valuable for creating a multi-dimensional understanding of the service benefits and can provide useful information to feed into Theory of Change models that usually incorporate stakeholder mapping and identifying pathways to change (Vogel, 2012).

### 5.2.10 Community opinions, awareness and attitudes and help-seeking sampling

This type of evidence seeks to explore broader social trends and capture effects at the community or national level. In relation to evidence of effectiveness, it can focus on three main areas – the attitude and awareness about child protection issues gathered in the community, for example, through periodical surveys; and population surveys seeking to establish the number of people who have contacted a helpline for help; or community surveys on public awareness of the helpline, for example, after awareness campaigns. An example of the latter is the evaluation of the publicity campaign (radio, press, outdoor and local PR) launched by the Scottish Government to raise awareness of the Child Protection Line (2008). The research aimed to assess awareness of the Child Protection Line; any change in public attitudes towards protecting vulnerable children; and likelihood of using the phone line. The study comprised pre- and post-campaign survey research conducted using the Scottish Opinion Survey omnibus (quota sampling) in comparison with two non-campaign areas. The study was able to demonstrate some improvement in outcomes (e.g., prompted awareness of the helpline), but in relation to most of the studied outcomes, the effects were small.

As there are no other examples of similar designs in the final sample, it is difficult to draw conclusions as to whether the findings indicate ineffectiveness of the campaign or methodological challenges. Still, the initial literature search revealed a considerably large body of health-based interventions aiming to alter social norms and practices that have been subject to thorough evaluations and were able to demonstrate effectiveness (these studies are not included in the final sample as the work of these helplines is not related to children's issues and does not usually involve contact related to a crisis situation). Our observation is also upheld by Wood and colleagues (2016), who suggest that there is a robust evidence base, including meta-analyses, on the effectiveness on interventions aiming to alter norms related to issues like smoking, drink-driving and alcohol abuse, but other issues like violence are seldom rigorously evaluated. We also found that when it comes to child-related services, the evidence base is even scarcer. Undoubtedly, the methodological challenges of attributing effects at the community and social level contribute to this (Coveney et al, 2012).

Another approach to using a population survey relates to establishing the number of people using a helpline. This method was adopted in the cross-national study of health-related help-seeking by Coulter (2006), who found that 28 per cent of the UK population had used a helpline to seek health advice, a significantly higher proportion than in any of the other countries (Australia, Canada, New Zealand, Germany and the US). While this study is able to compare public engagement

in relation to several indicators (quality communication; access to alternative sources of information and advice; support for self-care and self-management, etc.), this approach might struggle to draw evidence about the performance of a particular helpline in a context where several similar services operate. To cope with this challenge, a smaller-scale project might be more suitable. An example of such a study is one undertaken by Downing and Cook (2006) examining children's sexual health services in Knowsley, including the helpline Sexwise. Comparing the use of the helpline to other services available to children, the study used a questionnaire distributed in local schools, youth services and clinic services and focus groups with children. Downing and Cook (2006) found that the helpline Sexwise was one of the most frequently accessed information services: 26 per cent of the children who took part in the survey had used it (preceded only by a youth advisory centre – 34 per cent). They also refer to an earlier evaluation study that showed that the use of the helpline increased after the 'Sex Lottery' awareness campaign.

Once the challenges are overcome, there are substantial benefits of the population- or community-wide evidence-gathering methods – they might be particularly useful for establishing gaps in the service provision, comparing the effectiveness of different types of services or capturing the effects of particular campaigns.

### 5.2.11 Cost-benefit analysis

Cost-benefit analysis was used by five studies in the sample (Shelter, 2011; Starks et al, 2012; Bowles, 2014; Calderón et al, 2017; Cullen et al, 2017), and followed a similar approach of estimating the monetary value for the benefits being delivered by the helpline. The calculation of costs was fairly consistent and comprised an exhaustive list of the incurred expenditure, including recurrent costs for maintaining the services, as well as costs for set-up and sometimes roll-out. The approach to estimating the service benefits was more complex and the techniques varied based on the nature of the service provided. In his research on the child sexual abuse helplines Stop it now! UK and the Netherlands, Bowles (2014) reviewed several methods of estimating the benefit: valuing reductions in crime using estimates of the costs of an offence type as the basis or looking at the costs to victim/survivors and estimating factors like reduced lifetime earnings and a higher incidence of mental health issues later in life.

Calderón and colleagues (2017) take a different approach in their research on Extended HOPE, a helpline for out-of-hours support for children with mental health problems. They used the contact logs to calculate the avoided cost associated with prevented events, such as accidents and emergency visits, placement breakdowns, paediatric ward stays, etc. A possible simplified version of the cost-benefit

analysis is the calculation of the average recurrent costs per call that can be used as a measure of effectiveness, particularly in comparison with other helplines (see, for example, Bowles, 2014). While this analysis might be beneficial for established helpline services, newly set-up services will be facing a disadvantage as they will have a higher need for training and a less established social presence, which will affect their use. The third study, by Cullen and colleagues (2017), evaluated the programme 'Breaking the Cycle' that offers face-to-face support to families affected by parental substance misuse. This model used the number of referrals as the basis of calculating the cost-benefit, which requires knowledge of the outcome of the referrals that might be difficult to obtain in a helpline context for reasons of confidentiality and anonymity, as well as resource restrictions. If this data could be acquired via direct contact with the referral agencies, the model could offer a promising approach to cost-benefit analysis.

As the above-mentioned evaluation designs enjoy different popularity and weight in demonstrating service effectiveness, some gaps can be identified in the evidence terrain. In the discussion below, we focus on how these gaps can be approached, and suggest some strategies for overcoming the barriers based on the most compelling good practice examples identified by the review.

## 5.3 Overcoming barriers to effectiveness

- ***Multiple measures and indicators***: one of the strengths of the evaluations included in the sample relates to the use of mixed methods to demonstrate the effectiveness of the service. Given the sensitive settings in which child and youth helplines operate and the various channels of support that are used at present, it is positive that the development of evaluations often combines a range of methods, drawing on both quantitative and qualitative measures. The simplest designs bring together call log data with case studies or quotations from service users, thus capturing both the lived experience and the numerical evidence. The more complex designs rely on survey data and qualitative research, such as focus groups or interviews, as well as some support elements, such as evidence reviews and helpline output data. The use of multiple measures and indicators including models of quantifying qualitative or case study evidence can offer possible solutions to the quality of evidence barriers.
- ***Quality of evidence***: it might be beneficial to distinguish between 'hard' and 'soft' outcomes and to use different types of evidence to demonstrate effectiveness. Hard outcomes can be related to securing referrals and managing to resolve cases and might be quantified by attributing an economic value to the service. Soft outcomes can measure the work done to mitigate the impacts of

the difficulties that children experience and can be evaluated via outcome monitoring data or output data gathered by the local support services, as well as by qualitative data. Using a combination of some lower-quality data that a helpline can gather longitudinally and a single high-quality evaluation can offer an effective approach to demonstrating effectiveness.

- ***Using standardised measures***: the use of standardised measures with established reliability, validity and applicability to children is a good way of producing robust evidence of effectiveness. Using such measures will also facilitate longitudinal research designs and outcome comparability across the helpline/child support sector. There are, however, challenges related to using standardised measures associated with the suitability of the existing measures. First, the review demonstrated that there are no prevalent models of using established measures of effectiveness across the helpline sector. This makes identifying the best approaches more challenging. Second, the outcomes assessed by established standardised measures might not necessarily be the ones that represent most effectively the work of a particular helpline. It is important to find measures that capture the service benefits comprehensively. Third, the robust standardised measures are usually comprehensive and take a while to complete, which might make them unsuitable for a telephone environment, and the use of standardised measures requires a pre- and post-service evaluation, which might be too demanding of children, particularly younger ones. Nevertheless, we discussed a number of good examples that demonstrate that this can be done. Fourth, helplines that offer support in relation to a wider range of issues children experience, such as Childline, might struggle even more to find a concise measure that can be applied to all cases. The review demonstrated that helplines that have a more narrowly defined focus (e.g., suicide support) have more well-established measures used across different services. Using measures for subgroups of children based on the issue they experience might offer a way forward for helplines with a more general focus. Alternatively, a more wide-ranging approach that aims to capture and measure a combination of key (and more universal) outcomes might be suitable (see Table 3 below).

## Table 3: Mapping key areas of outcome evaluation (child- and counsellor-based evaluation) and suggested measures

| **Outcome (child evaluation)** | **Suggested measures (5-point scale, blank answer options, 3 emoticons:**  |
|---|---|
| Child wellbeing* | 5-point scale from "I feel much worse" to "I feel much better" |
| Perceived helpfulness* | 5-point scale from "It didn't help at all" to "It helped a lot" |
| Instrumental empowerment* | 5-point scale, After talking to [helpline] I have "… no idea about what to do" to "… an idea about what to do" |
| Self-confidence and emotional empowerment* | 5-point scale from "Disagree" to "Agree" |
| Satisfaction | 5-point scale from "Not satisfied at all" to "Completely satisfied" |
| Employee emotional support | 5-point scale from "The counsellor didn't care about me" to "… cared for me a lot" |
| Employee instrumental support | 5-point scale from "The counsellor gave me no information or advice" to "…lots of information and/or advice" |
| To be heard | 5-point scale from "Disagree" to "Agree" that the counsellor listened to what I wanted to say |
| Recommendation | 5-point scale from "Not likely at all" to "Extremely likely" that I will recommend the helpline to a friend or peer |
| Background variable | 4-point scale on previous contact with helpline: "No, this was my first time"; "Yes, a few times (1-3)"; "Yes, several times (4-10)"; "Yes, more than 10 times" |

| **Outcome (counsellor evaluation)** | **Measures (5-point scale, 1-5)** |
|---|---|
| Counsellor approach | Action-facilitating support: giving information, advice, tangible help, factual input and feedback on actions in order to assist the child in solving a problem<br><br>Emotional support: empathetic listening, expressing sympathy, care and concern, building up esteem without a direct effort to solve a problem |
| Satisfaction | 5-point scale from "Not satisfied at all" to "Completely satisfied" |
| Employee emotional support | 5-point scale from "Not at all" to "a lot" extent to which you provided emotional support |
| Employee instrumental support | 5-point scale from "Not at all" to "a lot" extent to which you provided informational support or advice |
| Self-confidence/ emotional empowerment | 5-point scale from "Disagree" to "Agree" at the end of the session the child had more belief in her/himself |

Note: *Outcome measure also used at follow-up evaluation.
Source: Based on van Dolen and Sindahl (2017b)

The model above, developed by Child Helpline International, offers a promising approach to measuring helpline outcomes based on child evaluation (both at the end of service and follow-up) and counsellor evaluation. More information about the measures is available in Sindahl (2013) and van Dolen and Weinberg (2017). The model combines items from existing scales that are already used in the context of helplines. These are planned as a collection of single-item measures (e.g., on satisfaction, empowerment, emotional support, etc.). Single-scale items are recommended for studies with children to increase likelihood of survey completion (Drolet & Morrison, 2001; Fukkink & Hermanns, 2009; van Dolen & Weinberg, 2017).

The review found a surprisingly small number of studies using validated measurements and scales. A number of studies in the sample used scales and measures (Reubsaet et al, 2006; Santor et al, 2007; Finn & Hughes, 2008; Karver et al, 2010; Phillips-Howard et al, 2010; Coveney et al, 2012; Wilson & Haner, 2012; Grove et al, 2015; Tyson et al, 2016; Donaldson & Irwin, 2017; Goldsmith, 2017; van Dolen & Sindahl, 2017b), but these were not standardised and there was no information on scale reliability and validity – hence, they are not included in the table below. For example, alongside standardised measures, Fukkink and Hermanns (2009) used a 7-item service satisfaction scale and a single-item perceived burden of the problem scale, but it is unclear if these measures have been validated. Similarly, Haner and Pepler (2017) use measures of distress, perceived difficulty and self-efficacy but do not discuss validation. While some of these studies demonstrate a promising approach to outcomes measures, the table below includes only validated standardised measures.

## Table 4: Examples of standardised measures used

| Outcome domain | Key measures | Key features | Source |
|---|---|---|---|
| Mental health and emotional symptoms | Children's Global Assessment Scale (CGAS) | General mental health functioning of children (under 18), score 1-100 | Calderón et al (2017) |
| | Clinical Global Impression (CGI) Scale | Severity of mental health illness 7-point scale, 3 items (symptom severity, treatment response and the efficacy of treatments) | Cullen et al (2017) |
| | Generalised Anxiety Disorder (GAD-7) | 7-item scale for clinical anxiety | Cullen et al (2017) |
| | Patient Health Questionnaire (PHQ-9) | 9-item depression severity measure | Cullen et al (2017) |
| | Shorter Warwick Edinburgh Mental Wellbeing Scale (SWEMWBS) | 7-item mental wellbeing scale, for children aged 10 or above | Cullen et al (2017) |
| | General Health Questionnaire (GHQ-12) | 12-item scale for non-psychotic and minor psychiatric disorders | King et al (2006) |
| | Beck Depression Inventory (BDI) | 21-item psychometric test for measuring the severity of depression | Rhee et al (2005) |
| | Brief Symptom Inventory (BSI) | 53-item self-report scale that evaluates psychological distress and psychiatric disorders | Rhee et al (2005) |
| | Brief Psychiatric Rating Scale (BPRS) | 24-item scale measuring psychiatric symptoms like anxiety, depression and psychoses, completed by the therapist | Rhee et al (2005) |
| | Visual Analogue Scale (VAS) | Clients visualise and then report their current level of depression on a scale from 0-100 | Rhee et al (2005) |
| Functional competence | Cantrill ladder | 1-item overall life satisfaction scale | Fukkink & Hermanns (2009) |
| | Youth Self-Report (YSR) syndromes | Scale for behavioural and emotional problems in children and adolescents over the age of 11; 8 items including anxious/depressed, withdrawn/depressed, somatic complaints, social problems, thought problems, attention problems, rule-breaking behaviour and aggressive behaviour | Haner & Pepler (2017) |

| Outcome domain | Key measures | Key features | Source |
|---|---|---|---|
| Multi-dimensional | Strengths and Difficulties Questionnaire (SDQ) | 25-item scale measuring conduct, hyperactivity, emotional, peer, pro-social factors | Fukkink & Hermanns (2009); Cullen et al (2017) |
| | Based on Perceived Effectiveness of Counselling Scale (PECS) | 16-item, 7-point scale measuring distress, problem difficulty, hope, problem clarity, self-efficacy | Law et al (2015) |
| Session-based measures | Telephone Crisis Support Skills Scale (TCSSS) | 23-item scale assessing workers' intentions to use recommended skills with callers | Kitchingman et al (2015) |
| | Session Impact Scale (SIS) | 16-item scale designed to measure therapeutic impact based on specific session content (task impacts, relationship impacts and hindering impacts) | King et al (2006) |
| | Therapeutic Alliance Scale (TAS) | 30-item scale measuring alliance between counsellor and child/adolescent (both a total scale score and subscales for mutual liking, resistance and collaboration) | King et al (2006) |
| | Therapeutic Alliance Quality Scale (TAQS) | 5 items on a 5-point scale on the working relationship between youth and therapist (bond and agreement on goals and tasks) | Hanley (2009) |
| | Crisis Call Outcome Rating Scale (CCORS) | 26-item scale with 16 items to indicate successful outcomes and 10 that indicate non-successful outcomes, rated on a 7-point scale (supportive approach and good contact; active listening; collaborative problem-solving; negative style) | Mishara et al (2007); Mokkenstorm et al (2017) |

- ***Mapping aims, outcomes, and measures***: having a clear model of how the organisational aims match the desired outcomes and how the evidence can support the demonstration of effectiveness is crucial. In the sample of evaluations, there are several good examples of mapping aims and outcomes and even more examples of mapping outcomes and measures. However, we did not find any comprehensive approaches that map aims, outcomes and measures. In addition, many of the evaluations did not include a clear justification of how the outcomes were chosen, and a clear strategy behind the selection was not apparent when reviewing the publications. In some cases, the methods used were also not clearly leading up to the claims of effectiveness, resulting in some overstatement of the effect of the interventions.

## Case study: National Sexual Assault Online Hotline (NSAOH), US

An example of mapping aims, outcomes and measures is the evaluation of the online hotline for rape and abuse, the National Sexual Assault Online Hotline (NSAOH), conducted by Finn and Hughes (2008). The service has three aims that are linked to desired and measurable outcomes. While this is not based on a Theory of Change model, the straightforward matching between the aims, outcomes and measures is helpful.

(1) Delivering a secure, anonymous, real-time online hotline: evaluation of this goal includes ease of use by visitors and volunteers, extent of programme disruption due to technical difficulties, examination of user waiting time and examination of volunteer self-scheduling process. (2) Providing effective crisis intervention, information and support: evaluated via user satisfaction and perceived helpfulness, user feedback and volunteer perceptions of programme usefulness. (3) Awareness and wide use of the service: number of visitors and volunteers and number of affiliations of the rape crisis centre.

This is a fairly simple model that can be improved by providing a justification of how the used outcomes were identified and why the measures were selected, including information about reliability and validity of the measures, as well as benefits in comparison to alternative measures.

# 6. Recommendations

- ***Building a model of effectiveness***: while the evidence base on child helpline effectiveness is growing and there are good evaluation examples, there are still no established models within the sector, and this area requires more efforts. Based on the existing examples and the evidence review findings, a model of effectiveness could include:

  1) *Context mapping*: contact analysis in relation to, for example, socioeconomic factors, policy and regulation; key actors, stakeholders and sources of influence; and child protection service provision gaps and opportunities for change.

  2) *Helpline planning*: description of aims, objectives, target population; activities and outputs (counselling; education and information; resources and referrals); human and material resources; procedures, protocols and professional development; and contributions from stakeholders, funders and donors.

  3) *Outcome planning*: description of desired outcomes (short-term, intermediate and long-term) from the work of the helpline at individual, interpersonal, organisational, inter-organisational and communal and social levels.

  4) *Pathways to effectiveness*: identifying ways to achieve the desired outcomes; criteria for effectiveness; and strategies to overcome barriers.

  5) *Monitoring and evaluation planning*: mapping the measures and evidence-gathering methods to the outcomes and criteria for effectiveness; and creating an evaluation-gathering strategy (methods, timelines, funding, etc.).

  6) *Monitoring and adaptation*: using the evidence to adapt the services, effectiveness model and future evaluations.

- ***Realistic and measurable outcomes at all levels***: service evaluations need to be based on realistic expectations of what is achievable for child helpline services and an in-depth understanding of the limitations to the positive effects, particularly in relation to more vulnerable groups of children or effects occurring at community or social levels. The planned outcomes need to be mapped to associated measures and evaluation strategies in order to demonstrate effectiveness. This is more easily achievable in relation to individual outcomes, while good evidence-gathering examples related to outcomes on an inter-personal, organisational, community or national level are scarcer and more inconsistent. These gaps demonstrate the need for further efforts in this area. However, the existing gaps also reflect the nature of helpline support – the brief, anonymous, and often singular contact with the

callers, the sensitivity of the issues discussed, and the vulnerability of some of the children. This makes evaluations particularly challenging, both ethically and in practice, suggesting the need for service-specific criteria and standards for evaluation that have a greater emphasis on less intrusive and more child-friendly and inclusive methods.

- ***Establishing quality standards***: the lack of established and recognised evaluation models and outcomes measures within the sector impedes the ability of child helplines, both individually and collectively, to make claims about the effectiveness of their service. While some examples exist, most of the evidence does not follow established evaluation standards for interventions (for reasons that include the nature of the work). The difficulties in obtaining baseline measures, comparison groups and follow-up data pose key challenges to demonstrating effectiveness. While these challenges are substantial for individual helplines to tackle, some collaborative work on exchanging knowledge and good practice could help to improve practice and comparability by identifying models applicable across child helplines and local contexts. Examples of good practice in this area exist, such as the models of Child Helpline International's collaborative international work and Insafe's helpline networks (Dinh et al, 2016; Child Helpline International and Oak Foundation, 2017).
- ***Comprehensive service evaluations***: it is common for interventions to seek multiple channels for engagement, taking advantage of the opportunities provided by the digital environment (Cronin et al, 2017). Therefore, it is more helpful to perceive child helplines as a complex hybrid of digitally mediated support and to seek evaluation models that can capture effectively this multiple character. The fast adoption of digital technologies for child support purposes seems to precede an informed and evidence-based understanding of what works for child support, and how online and mobile services can supplement the existing portfolio of services most effectively. More comparative evidence is needed to draw transferrable lessons about the most effective models of using multiple channels for child support. Hence, service evaluations should try to distinguish between different service channels in terms of users, outcomes and effectiveness.
- ***Role of child helplines in systemic change***: funding larger-scale evaluations that can explore the relationship between children's needs, the available support and the effective coordination of child support agencies can help to diagnose gaps and inform better child protection efforts. The evidence and evaluation are crucial to child protection systems' understanding and service provision.

- ***Building strategic and inclusive child support alliances*** seems crucial in overcoming not only the evidence gaps and methodological challenges, but also in disseminating good practice and working collaboratively to improve child support and protection services. The digital environment has presented new challenges to child protection, but it has also afforded new opportunities for engaging children more actively in voicing their concerns, offering peer support and designing more child-inclusive policy and practice. Gathering a more robust evidence base drawing on children's own perspectives of their needs, their experiences of service provision and their evaluation of the remaining gaps is an important step in this direction.

# Appendix 1: Definitions

We define a number of key terms to clarify how they have been used in the review. These might differ from how the terms have been used in the original studies we reviewed, but establishing a unified language was important for the consistency of the report.

## Child

Following the United Nations (UN) Convention on the Rights of the Child, we define a 'child' as a person under the age of 18. However, the different services covered in the review have target groups of varying ages. The information on the age group of each service can be found in the report supplement (organised by the author of the study).

## Helpline

'Helpline' describes a service that provides listening and emotional support as well as information to assist users in dealing with issues they face in their everyday lives (Dinh et al, 2016). Traditionally, helplines offer support primarily via telephone, but new technologies have created different channels for communication (Dinh et al, 2016). Hence, when referring to 'helpline' we include the full range of different channels used to deliver a particular service, such as chat counselling, text message or email support, online forums or apps. This approach is adopted as the review demonstrated that most helplines function on various platforms. At the same time, there is not sufficient evidence yet to explore the outcomes via channel, even though a number of studies exist.

Throughout the report we also use 'service' and 'intervention' to refer to helplines. While we acknowledge that these three terms are generally not synonymous, we use them interchangeably to facilitate the flow of the argument.

## Evaluation

This relates to the relevance, effectiveness, efficiency, impact and sustainability of a project or programme (Child Helpline International and Oak Foundation, 2017). There is variation in the definition of 'evaluation' in the social research literature. In some of the literature reviewed here, the term 'evaluation' is used to describe an overview of the activities of the helpline, the profile of its users and the types of issues addressed. In other cases, it refers to the analysis of changes in user behaviour, knowledge or perception as a result of the helpline actions. In yet other cases, it describes comprehensive studies that

seek to establish the multiple impacts of the helpline starting from the individual and scaling up to the social level.

## Impact

'Impact' can be understood to mean the more general effects of service provision on the sector or society as a whole (University of Central Lancashire and NCB Research Centre, 2016).

## Outcome

'Outcome' is the short- and medium-term effect of an intervention's outputs, such as change in knowledge, attitudes, beliefs or behaviours.

## Output

'Outputs' are the results of the child helpline activities; the direct products or deliverables of the child helpline's activities, such as the number of counsellor training sessions completed; the number of children served; or the number of calls answered. In some studies, this is referred to as the 'process' (see, for example, Abba, 2001).

## Referral

A 'referral' is defined as "a request for services to be provided by children's social care and is in respect of a child who is not currently in need. A referral may result in: an assessment of the child's need; the provision of information or advice; referral to another agency; or no further action" (DfE, 2017a: p.5).

## Systematic evidence mapping

'Systematic evidence mapping' refers to a review process that systematically identifies and describes the research that has been undertaken within the boundaries of the review question (EPPI-Centre, 2018). For further details, see Appendix 2.

## Theory of Change

'Theory of change' is "an outcomes-based approach that applies critical thinking to the design, implementation and evaluation of initiatives and programmes intended to support change in their contexts" (Vogel, 2012: p.3).

# Appendix 2: Detailed methodology

## Approach

The aim of the review was to systematically map the existing literature on the effectiveness of child helplines with a particular focus on identifying how outcomes were defined, exploring the range of possible approaches to measuring the effectiveness of children's helplines and evaluating the strength of the evidence. We predicted that the body of literature focusing directly on the effectiveness of helplines would be relatively sparse, hence the team adopted an inclusive approach where a comprehensive literature search would be carried out using a broader range of search words. The following inclusion criteria was applied:

- The review included evidence relating to child helplines (expanded with good examples from adult helplines, as relevant, particularly related to supporting children indirectly via working with adults, such as parents or educators).
- Evidence relating to both shorter- and longer-term outcomes from service use including outcomes at an individual, interpersonal, organisational (internal or external), community or social level.
- The review covered a range of services including counselling interventions and support delivered by telephone, online, via digital devices or apps.
- There were no limits on geographical scope, but the review only included publications available in English.
- The review focused on work published since 2000 (for relevance to the current context).
- Peer-reviewed research outputs were prioritised for quality control but publications from non-governmental organisations (NGOs), government reports, industry sources and other relevant grey literature is also included (providing the sources meet quality requirements).
- The review drew on high-quality and methodologically reliable research – both in terms of the systematic mapping of the evidence and also the sources included in the analysis.

A systematic evidence review approach was considered as this is the dominant method for synthesising the evidence on the effectiveness of health and social interventions (Gough et al, 2012). However, we found that the robust application of the requirement for statistical meta-analysis of controlled trials, which is at the core of systematic evidence reviews (Gough et al, 2012), would be too limiting and would substantially reduce the evidence base the review could draw on. The team intentionally sought an approach that would be better able to address the complexity of child helpline interventions and allow for understanding the research evidence within its social and paradigmatic context. Therefore, we applied a systematic mapping of the evidence that allowed us to describe the nature of the research field; facilitated the interpretation of the findings; and informed the final synthesis of the findings (Gough et al, 2012).

The systematic mapping comprised a more inclusive search strategy that allowed the inclusion of a broader range of sources, such as end-of-year reports, policy recommendations, methodological guides and case studies. To ensure maximum insight and rigour, the team also requested input from a range of experts on recommended research sources. A list of suggested experts was drafted at the beginning of the project and gradually expanded via snowballing and recommendations. The resources suggested by the experts were included in the evidence review database and substantially contributed to the overall rigour of the review.

## Search terms and outcomes

After consultation with the LSE subject librarian Heather Dawson, we selected 18 databases based on their suitability to the review scope and aims (see below). These cover the areas of general social sciences, health, government publications and grey literature.

The team compiled a combination of search words that fell into three groups: help and service terms, outcome terms, and child terms. Search testing was carried out to ensure optimal coverage of the review and efficient combination of the search terms (details available on request). The initial searches were conducted using the following combination of terms:

Group 1, help and service terms: *help* OR support* OR advice OR protect* OR safe OR hotline OR counsel* OR service OR interven**

Group 2, outcome terms: *outcome OR impact OR evaluat* OR effect**

Group 3, child terms: *child* OR young OR youth OR teen* OR adolescen* OR minor OR kid OR girl OR boy OR student*

Following some fine-tuning of the search words, the final search combination included:

*(1) helpline* AND (2) outcome* OR evaluat* AND (3) Group 3 (child words) in title OR abstract OR keyword* (where keyword search was available).

This search produced a total of 1,198 resources from all the databases. In some cases, when grey literature or government databases produced a very small number of results, a slightly broader search approach was adopted to boost the representation of these databases, as below. However, a broader search was not possible for all databases. For example, only the Web of Science database produced 24,639 results with Group 1, 2 and 3 words combined, and the database contained a large proportion of resources irrelevant to the scope of the review.

## Table 5: Databases, search protocol and results

| Database | Search words | Period | Search areas | Language filter applied | Results |
|---|---|---|---|---|---|
| Web of Science | helpline & outcome OR evaluat* & Phase 3 (child) | 2000-18 | Topic (title, abstract, key words) | English | 120 |
| Scopus | helpline & outcome OR evaluat* & Phase 3 (child) | 2000-18 | Title, abstract, key words | English | 100 |
| International Bibliography of the Social Sciences (IBSS) | helpline & outcome OR evaluat* & Phase 3 (child) | 2000-18 | Title or abstract | English | 5 |
| PAIS International | helpline & outcome OR evaluat* & Phase 3 (child) | 2000-18 | Title or abstract | English | 0 |
| PsycINFO (via Ovid) | helpline & outcome OR evaluat* & Phase 3 (child) | 2000-18 | Title or abstract | English | 19 |
| EMBASE (via Ovid) | helpline & outcome OR evaluat* & Phase 3 (child) | 2000-18 | Title or abstract | English | 61 |
| Cinahl (via Ebsco) | helpline & outcome OR evaluat* & Phase 3 (child) | 2000-18 | Title or abstract; child in abstract only | English | 67 |
| MEDLINE via Ovid | helpline & outcome OR evaluat* & Phase 3 (child) | 2000-18 | Title or abstract; child in abstract only | English | 31 |

| Database | Search words | Period | Search areas | Language filter applied | Results |
|---|---|---|---|---|---|
| HMIC database | helpline & outcome OR evaluat* & Phase 3 (child) | 2000-18 | Title or abstract; child in abstract only | English | 2 |
| Child Development & Adolescent Studies (via Ebsco) | helpline & outcome OR evaluat* & Phase 3 (child) | 2000-18 | Title or abstract | English | 7 |
| Criminal Justice Abstracts (via Ebsco) | helpline & outcome OR evaluat* & Phase 3 (child) | 2000-18 | Title or abstract | English | 4 |
| SocINDEX (via Ebsco) | helpline* & outcome* OR evaluat* & Phase 3 (child) | 2000-18 | Title or abstract | English | 6 |
| Public Information Online | helpline* | 2000-18 | Title | N/A | 2 |
| NHS evidence | child* helpline evaluat* | 2000-18 | All (no options available) | N/A | 763 |
| Social Care Online | Phase 1 (help terms) outcome* or evaluat* | N/A | Title or abstract | N/A | 5 |
| Open Grey | helpline & evaluat* | 2000-18 | N/A | English | 2 |
| SSRN Papers | helpline* | all | Title, abstract, key words | N/A | 4 |
| NSPCC | Evaluat* OR outcome* OR effect* OR review* AND counselling service*; helpline* OR hotline* AND evaluat* OR outcome* OR effect*; treatment OR intervention AND counselling service* | All | A combination of key words and all content | N/A | 197 |
| Expert recommendations | N/A | 2000-18 | N/A | N/A | 26 |
| TOTAL search results | | | | | **1,421** |

## Databases searched

### General social sciences

- Web of Science – provides access to articles covering all aspects of the sciences, social sciences and humanities.
- Scopus – covers a wide range of science and social science subject areas including gender studies, women's studies and LGBT issues.
- International Bibliography of the Social Sciences (IBSS) – includes over 3 million bibliographic references dating back to 1951.
- PAIS International – produced by the Public Affairs Information Service, this indexes the content (often with abstracts) of over 1,000 journals as well as some books, theses and government documents, on the subjects of public affairs, international relations, social policy and other social science subjects. Coverage is from 1972 onwards.
- Child Development & Adolescent Studies (via Ebsco) – this bibliographic database is today's source for references to the current and historical literature related to the growth and development of children through to the age of 21.
- Criminal Justice Abstracts (via Ebsco) – this bibliographic database provides records selected from the most notable sources in the criminal justice field. It covers journals from around the world, reflecting the increasing globalisation of criminology studies.
- SocINDEX (via Ebsco) – this bibliographic database provides high-quality indexing and abstracts for journals covering the broad spectrum of sociological study. It offers millions of records with subject headings from a sociological thesaurus.

### Health databases

- PsycINFO – via Ovid, a specialist database of the American Psychological Association, providing abstracts of articles relevant to all fields of psychology from the 19th century to the present day.
- EMBASE – via Ovid, indexes medical, biomedical and neuroscience journal articles published since 1947.
- Cinahl – via Ebsco, specialist nursing and medical care subject index, referencing journal articles and reports published since 1937.
- MEDLINE via Ovid (subscription version of PubMed).
- HMIC database – via Ovid, Health Management Information Consortium, contains references to reports, articles and books from the King's Fund Library and Department of Health databases. Covers health policy, health services and healthcare, focusing mainly on the UK since 1979.

### UK government publications

- Public Information Online – the full text of UK government documents and reports since 2006.
- NHS evidence – free website searches on NHS guidance, plus journal articles (content restricted).

### Grey literature

- NSPCC library – the only UK library dedicated to child protection, child abuse and child neglect, holding over 40,000 resources including inquiry reports and case reviews, training resources and practice toolkits, international journals and grey literature.
- Social Care Online – the UK's largest database of information and research on all aspects of social care and social work.
- Open Grey – system for information on grey literature in Europe, with open access to 700,000 bibliographical references of grey literature (paper) produced in Europe.
- SSRN Papers – research repository that spans across multiple disciplines.

### Expert literature

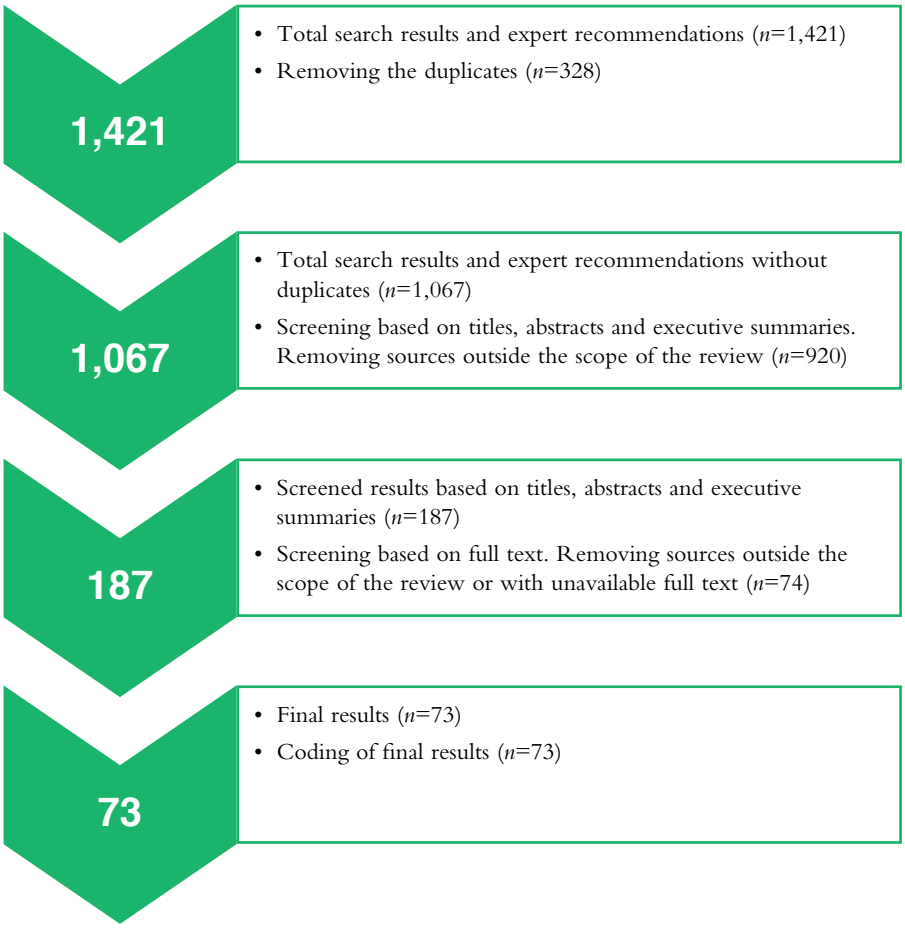
- We included 26 sources recommended by experts in the field.

## Screening

First, the batch of 1,421 studies ('search results') was cleaned by removing duplicates. This reduced the search results to 1,067 entries, which were screened for relevance. The screening of these studies proceeded through two stages.

At the first stage, the results were screened on the basis of titles, abstracts and executive summaries to judge their relevance to the inclusion criteria. This produced a database of 187 results ('screened results').

The second stage of screening was done on the basis of full text, applying the same criteria for relevance and a new requirement for methodological rigour. This produced a final batch of 73 results that were coded for the analysis.




## Screening criteria

The first stage of screening involved removing non-publication results (e.g., videos mainly from the NSPCC database) and sources published earlier than 2000 (for databases where filtering on the basis of year of publication and resource type was not available). Most of the results excluded at this stage pertained to adults (most often related to smoking cessation, AIDS or cancer support, alcohol use, gambling, breathing problems and influenza).

The presence of these sources can be explained by the use of the word 'young' in the search terms that can relate to older and younger groups of adults. Another group of excluded publications on adult-based services related to parent-targeted services, including helplines. These were often related to issues like reproduction, childbirth, breastfeeding and children's health (e.g., diabetes, allergies, skin conditions, etc.) or behaviour (infant crying, anti-social behaviour). A number of search results (particularly in the NHS database) related to guides on professional practice or quality standards, handbooks on service delivery, information provision for patients, government policies, national plans or strategy documents or evaluations of other services (e.g., healthcare provision), which recommended the use of a helpline but did not deal with helpline evaluation.

During the first stage screening, a small number of adult-related publications remained included in the sample due to their perceived methodological relevance. Finally, a number of results discussed a range of child services but did not include any evaluation or evidence of effectiveness. They were removed as they could not contribute to the aims of the review.

At the second stage, the full text of the screened results was examined, reviewing the sections on methodology and study design of each full text and removing sources that were irrelevant (did not include service evaluation or any evidence on service outcomes) or which were drawing on unreliable methodology. Reviewing the full text allowed a more refined decision about relevance, which was not possible based on abstract and title only. Resources where the full text was not available were also removed at this stage. This included a number of conference abstracts for which a full-text publication was unavailable (mainly EMBASE database).

## Coding

The final results were coded via a coding template (see Appendix 3) that was developed in advance as part of the research proposal methodology. The coding template was created based on a quick preliminary literature scoping and was later updated during the systematic evidence mapping to ensure that it captured all that was relevant for the review information. The coding involved noting down key information about the service, the study and the findings.

The *service data* involved coding the publication in terms of target group, channel (telephone support, online chat, text message service, etc.), availability (when the service was accessible), type of user (user data including demographic characteristics), geographic location of the service, as well as a short description of the service offered.

The coding of the *type of study* included a classification of the study (conceptual; literature review/systematic review; primary research; secondary analysis), the methodology (experimental study, quasi-experimental/ex post facto, exploratory, case study, etc.) and type of data used (qualitative, quantitative, mixed methods). Information about the evaluation method, including sample and age group, recruitment and ethics, were also recorded (when available). The coding frame also contained a description of the research questions or aims and comments about the study value and reliability, or limitations in relation to the scope of the evidence review. Here, brief comments on the value of the study and place in relation to other studies were made, and it was noted whether the research engaged in a discussion of outcomes and their measurement. The comments related to framing, gaps or critique were not necessarily related to the value of

the study itself but reflected the extent to which it contributed to the current review.

Finally, the coding of the *study findings* offered a brief comment on claims and findings that might be useful for child helplines (rather than a full report of all the findings). It also contained a description of how the study conceptualised the outcomes of the intervention. In cases where the text did not define clearly the service outcomes but there was some relevant discussion, this was also included. Finally, the coding referred to information of how the study identified its effectiveness. As for outcomes, not all studies focused on effectiveness or included a discussion of how their outcomes were measured. In some cases, understanding of effectiveness was implied in the text and efforts were made to capture this in the coding framework. Some of the studies offered additional useful information, for example, case study material. This was also added, if deemed relevant to the review.

The coded studies are available as a review supplement.

# Appendix 3: Coding frame

| Area | Code | Explanation |
|---|---|---|
| Service data | Type of service | Description of the service offered:<br>• Target group: adult/child helpline<br>• Channel: counselling, telephone support, online support, support via digital devices and apps<br>• Availability: time |
| | Type of user | User data including demographic characteristics |
| | Place/geographic location | Describing the location of the services |
| Study data | Type of study | • Conceptual<br>• Literature review/systematic review<br>• Primary research<br>• Secondary analysis |
| | Methodology and type of data | • Data: qualitative, quantitative, mixed methods<br>• Methodology: experimental study, quasi-experimental/ex post facto, exploratory, case study, other<br>• Evaluation method, including sample and age group |
| | Research themes or questions | Description of the research questions or aims |
| | Study value and reliability, limitations | Brief comment on value of study and place in relation to other studies; comments related to epistemology/framing/gaps or critique |
| Findings data | Key findings useful for child helplines | Brief comment on claims and findings |
| | Effectiveness | Description of how effectiveness has been evaluated, for example:<br>• Awareness: knowledge of available services, ways to access<br>• Channel: accessibility, approachability, appropriateness, appeal<br>• Message: clarity, usefulness, correct amount<br>• Competence: sincerity, respect, attention, empathy, supportiveness, reassurance, expertise<br>• Dropout |
| | Type of outcome | • Level: individual, interpersonal, organisational (internal or external), community or social<br>• Immediate/intermediate/final or longer-term |
| | Outcomes details | Description of the established outcomes, for example:<br>• Expectations met<br>• User satisfaction<br>• Perceived burden of the problem<br>• User wellbeing (e.g., improved performance)<br>• Further service use/referral, etc. |

# Appendix 4: Coded sources


Abba, K. (2001) *A review of drug prevention work and its evaluation in Stockport*. Liverpool: Public Health Institute, Liverpool John Moores University.

Andersson, K. & Osvaldsson, K. (2011) *Evaluation of BRIS' internet-based support contacts*. Linköping: Linköping University.

Bambling, M., King, R., Reid, W. & Wegner, K. (2008) Online counselling: The experience of counsellors providing synchronous single-session counselling to young people. *Counselling & Psychotherapy Research*, 8, 2, pp.110–116.

Boddy, J. & Smith, M. (2007) *Evaluation of the Parentline Plus helpline: Key findings*. London: Thomas Coram Research Unit.

Bowles, R. (2014) *Economic analysis of the Stop it Now! UK and Stop it Now! Netherlands helplines: Summary*. London: National Centre for Social Research (NatCen).

Burns, J. & Birrell, E. (2014) Enhancing early engagement with mental health services by young people. *Psychology Research and Behavior Management,* 7, pp.303–312.

Calderón, A., Edbrooke-Childs, J., Chapman, L. et al. (2017) *Extended HOPE service: Evaluation report*. London: Department for Education (DfE).

Child Helpline International and Oak Foundation (2017) *Child impact assessment manual: Assessing your child helpline's impact on children*. Child Helpline International.

Child Helpline International and UNICEF (2017) *A new reality: Child helplines report on online child sexual exploitation and abuse from around the world*. Child Helpline International.

Childline (2012) *Saying the unsayable: What's affecting children in 2012*. London: NSPCC.

Childline (2016) *It turned out someone did care: Childline annual review 2015–16*. London: NSPCC.

Cossar, J., Brandon, M., Bailey, S., Belderson, P. & Biggart, L. (2013) *'It takes a lot to build trust': Recognition and telling: Developing earlier routes to help for children and young people*. London: Office of the Children's Commissioner.

Coulter, A. (2006) *Engaging patients in their healthcare: How is the UK doing relative to other countries?*. Oxford: Picker Institute Europe.

Coveney, C.M., Pollock, K., Armstrong, S. & Moore, J. (2012) Callers' experiences of contacting a national suicide prevention helpline: Report of an online survey. *Crisis: The Journal of Crisis Intervention and Suicide Prevention,* 33, pp.313–324.

Cronin, C., Sood, S. & Thomas, D. (2017) From innovation to transcreation: Adapting digital technologies to address violence against children. *Child Abuse Review,* 26, pp.215–229.

Cullen, M.A., Simpson, C., Bubley, R., Pike, M., Cullen, S. & Allen, B. (2017) *Delivering family interventions through social investment? Sharing the learning from a voluntary sector case study*. Addaction.

Daly, M., Bray, R., Bruckauf, Z. et al. (2015) *Family and parenting support: Policy and provision in a global context*. Florence: UNICEF Office of Research – Innocenti.

Danby, S. & Emmison, M. (2012) Kids, counsellors and troubles-telling: Morality-in-action in talk on an Australian children's helpline. *Journal of Applied Linguistics and Professional Practice,* 9, pp.263–285.

Dinh, T., Farrugia, L., O'Neill, B., Vandoninck, S. & Velicu, A. (2016) *Insafe helplines: Operations, effectiveness and emerging issues for internet safety helplines*. European Schoolnet (EUN), Insafe and the Kaspersky Helpline Fund.

Donaldson, R. & Irwin, J. (2017) *YoungMinds: Parents Helpline evaluation report*. Fiveways.

Dottridge, M. (2008) *Kids abroad: Ignore them, abuse them or protect them?* Geneva: Terre des Hommes.

Downing, J. and Cook, P.A. (2006) *Evaluation of young people's contraceptive and sexual health services in Knowsley*, Liverpool: Public Health Institute, Liverpool John Moores University.

Duerr Evaluation Resources (2016) Evaluation report for iCanHelpline: A social media helpline for schools. Unpublished report. Duerr Evaluation Resources.

Emmison, M. & Danby, S. (2007) *Troubles announcements and reasons for calling: Initial actions in opening sequences in calls to a national children's helpline. Research on Language and Social Interaction,* 40, 1, pp.63–87.

Fernandes, P. (2006) *Evaluating children's participation in decision making: A case study of a child helpline in India*. London: Department of Social Policy London School of Economics and Political Science.

Finn, J. & Hughes, P. (2008) Evaluation of the RAINN National Sexual Assault Online Hotline. *Journal of Technology in Human Services,* 26, pp.203–222.

Fry, D., Casey, T. & NSPCC (2017) *Bringing the global to the local: Review of global trends in the prevalence and services for child maltreatment in order to inform research, policy and practice in England*. London: Early Intervention Foundation.

Fukkink, R.G. & Hermanns, J.M.A. (2009) Children's experiences with chat support and telephone support. *Journal of Child Psychology and Psychiatry,* 50, pp.759–766.

Gallagher, M. (2013) *Helpline highlight: Northern Ireland: A year in review: April 2011–March 2012*. London: NSPCC.

Gallagher, M. (2014) *Speaking up for Scotland's children: The NSPCC's helpline in Scotland: A year in review: April 2012–March 2013*. Edinburgh: NSPCC.

Goldsmith, C. (2017) *At the end of the line: Independent evaluation of the Pact Helpline*. North Research, Training and Development.

Graugaard, C., Nielsen, V.M.L, Jensen, S.O. et al. (2017) What do Danish youth want to know about sex? Twenty years of telephone counselling on sexual and reproductive issues. *Journal of Sex & Marital Therapy,* 43, pp.159–168.

Grove, C., Reupert, A. & Maybery, D. (2015) Peer connections as an intervention with children of families where a parent has a mental illness: Moving towards an understanding of the processes of change. *Children and Youth Services Review,* 48, pp.177–185.

Haner, D. & Pepler, D. (2017) Adolescents show positive changes in distress and hope after single session, post-based, anonymous counselling at Kids Help Phone. *Children and Youth Services Review,* 82, pp.207–213.

Hanley, T. (2009) The working alliance in online therapy with young people: Preliminary findings. *British Journal of Guidance & Counselling,* 37, 3, pp.257–269.

Hepburn, A. (2005) 'You're not takin' me seriously': Ethics and asymmetry in calls to a child protection helpline. *Journal of Constructivist Psychology,* 18, 3, pp.253–274.

Hudson, B. (2005) 'Not a cigarette paper between us': Integrated inspection of children's services in England. *Social Policy and Administration,* 39, pp.513–527.

Hutchinson, D. (2012) *Helpline highlight: A year in review*. London: NSPCC.

Hutson, N. & Cowie, H. (2007) Setting up an email peer support scheme. *Pastoral Care in Education,* 25, 4, pp.12–16.

Karver, M.S., Tarquini, S.J. & Caporino, N.E. (2010) The judgment of future suicide-related behavior: Helpline counselors' accuracy and agreement. *Crisis: The Journal of Crisis Intervention and Suicide Prevention,* 31, pp.272–280.

King, R., Bambling, M., Reid, W. & Thomas, I. (2006) Telephone and online counselling for young people: a naturalistic comparison of session outcome, session and impact therapeutic alliance. *Counselling & Psychotherapy Research,* 6, 3, pp.175–181.

Kitchingman, T.A., Wilson, C.J., Caputi, P., Woodward, A. & Hunt, T. (2015) Development and evaluation of the Telephone Crisis Support Skills Scale. *Crisis: The Journal of Crisis Intervention and Suicide Prevention,* 36, pp.407–415.

Law, J., Haner, D., Simon, A. et al. (2015) *Proof positive: Evaluation of Kids Help Phone's phone and live chat counselling services*. Kids Help Phone.

Mishara, B.L., Chagnon, F., Daigle, M. et al. (2007) Which helper behaviors and intervention styles are related to better short-term outcomes in telephone crisis intervention? Results from a silent monitoring study of calls to the US1-800-SUICIDE Network. *Suicide and Life-Threatening Behavior,* 37, pp.308–321.

Mokkenstorm, J.K., Eikelenboom, M., Huisman, A. et al. (2017) Evaluation of the 113Online suicide prevention crisis chat service: Outcomes, helper behaviors and comparison to telephone hotlines. *Suicide and Life-Threatening Behavior,* 47, pp.282–296.

Morgan, K., Chakkalackal, L. & Cyhlarova, E. (2012) *Life lines: Evaluation of mental health helplines*. London: Mental Health Foundation.

Nicholas, L. & Broadbent, S. (2015) *Collective intelligence in patient organisations*. London: National Endowment for Science, Technology and the Arts (NESTA).

NSPCC (2016) *Childline: 30 years of listening to children*. London: NSPCC.

NSPCC (2016) *Getting help: What children tell us about accessing services after sexual abuse*. London: NSPCC.

Ofsted (2004) *Every Child Matters: Inspection of children's services; Key judgements and evidence*. London: Ofsted.

Phillips-Howard, P., Madden, H., Mason, L. et al. (2010) *The WiSHing Project: Wirral Sexual Health, health and wellbeing needs assessment of young people*. Liverpool: Public Health Institute, Liverpool John Moores University.

Potter, J. & Hepburn, A. (2005) Discursive psychology as a qualitative approach for analysing interaction in medical settings. *Medical Education,* 39, pp.338–344.

Potter, J. & Hepburn, A. (2012) Somewhere between evil and normal: Traces of morality in a child-protection helpline. *Journal of Applied Linguistics and Professional Practice,* 9, pp.245–262.

Rantzen, E. & NSPCC (2017) *Not alone anymore: Childline annual review 2016/17*. London: NSPCC.

Reubsaet, A., Lechner, L. & de Vries, H. (2006) The Dutch cancer information helpline: More critical patients after 10 years. *Patient Education & Counseling,* 63, pp.215–222.

Rhee, W.K., Merbaum, M., Strube, M.J. & Self, S.M. (2005) Efficacy of brief telephone psychotherapy with callers to a suicide hotline. *Suicide & Life-threatening Behavior,* 35, 3, pp.317–328.

Santor, A., Poulin, C., Leblanc, J. & Kusumakar, V. (2007) Online health promotion, early identification of difficulties, and help seeking in young people. *Journal of the American Academy of Child & Adolescent Psychiatry,* 46, pp.50–59.

Scottish Government (2008) *Child protection line campaign evaluation: Report for the Scottish Government and the Union*. Edinburgh: TNS System Three.

Sharratt, K., Porter, J. & Truman, C. (2014) 'Nowhere else to turn' – Key findings from an evaluation of the offenders' families helpline. *Prison Service Journal*, pp.30–35.

Shelter (2011) *Keys to the future: An independent evaluation of Shelter Children's Services: England and Scotland*. Shelter.

Shor, R. & Birnbaum, M. (2012) Meeting unmet needs of families of persons with mental illness: Evaluation of a family peer support helpline. *Community Mental Health Journal,* 48, pp.482–488.

Sindahl, T.N. (2013) *Chat counselling for children and youth – A handbook*. Child Helpline International.

Starks, L., Kendall, S. & Whitley, J. (2012) *Reach: A new model of intervention for children before, during and after they run away: Executive summary*. Sandbach: Railway Children.

Statham, J. & Carlisle, J. (2004) *Calling for help: The role of telephone helplines in child protection. A study for the Children's Safeguards Unit*. London: Thomas Coram Research Unit.

Tyson, P., Law, C., Reed, S., Johnsey, E., Aruna, O. & Hall, S. (2016) Preventing suicide and self-harm. *Crisis: The Journal of Crisis Intervention & Suicide Prevention,* 37, pp.353–360.

University of Central Lancashire & NCB (National Children's Bureau) Research Centre (2016) *Independent advocacy: Impact and outcomes for children and young people*. Office of the Children's Commissioner.

van Dolen, W. & Sindahl, T.N. (2017) *Impact assessment: A guide for child helplines*. Amsterdam: Child Helpline International.

van Dolen, W. & Sindahl, T.N. (2017) *Impact report: Child helplines*. Amsterdam: Child Helpline International.

van Horn, J., Eisenberg, M., Nicholls, C.M. et al. (2015) Stop it Now! A pilot study into the limits and benefits of a free helpline preventing child sexual abuse. *Journal of Child Sexual Abuse,* 24, pp.853–872.

WHO (World Health Organization) (2009) *Changing cultural and social norms that support violence*. Geneva and Liverpool: WHO and Public Health Institute, Liverpool John Moores University.

Wilson, T. & Haner, D. (2012) *Live chat counselling pilot*. Kids Help Phone.

Wood, M., Wanless, P. & NSPCC (2016) *Safer childhoods, brighter futures: Keeping children safe today and for generations to come: Annual report and accounts 2015/16*. London: NSPCC.

Ybarra, M.L. & Eaton, W.W. (2005) Internet-based mental health interventions. *Mental Health Services Research,* 7, 2, pp.75–87.

# References


Abba, K. (2001) *A review of drug prevention work and its evaluation in Stockport*. Liverpool: Public Health Institute, Liverpool John Moores University.

Andersson, K. & Osvaldsson, K. (2011) *Evaluation of BRIS' internet-based support contacts*. Linköping: Linköping University.

Bambling, M., King, R., Reid, W. & Wegner, K. (2008) Online counselling: The experience of counsellors providing synchronous single-session counselling to young people. *Counselling & Psychotherapy Research,* 8, 2, pp.110–116.

Bowles, R. (2014) *Economic analysis of the Stop it Now! UK and Stop it Now! Netherlands helplines: Summary*. London: National Centre for Social Research (NatCen).

Burns, J. & Birrell, E. (2014) Enhancing early engagement with mental health services by young people. *Psychology Research and Behavior Management,* 7, pp.303–312.

Calderón, A., Edbrooke-Childs, J., Chapman, L. et al. (2017) *Extended HOPE service: Evaluation report*. London: Department for Education (DfE).

Child Helpline International and Oak Foundation (2017) *Child impact assessment manual: Assessing your child helpline's impact on children*. Child Helpline International.

Child Helpline International and UNICEF (2017) *A new reality: Child helplines report on online child sexual exploitation and abuse from around the world*. Child Helpline International.

Childline (2012) *Saying the unsayable: What's affecting children in 2012*. London: NSPCC.

Childline (2016) *It turned out someone did care: Childline annual review 2015–16*. London: NSPCC.

Childline India (2018) *About us*. www.childlineindia.org.in/1098/1098.htm

Cossar, J., Brandon, M., Bailey, S., Belderson, P. & Biggart, L. (2013) *'It takes a lot to build trust': Recognition and telling: Developing earlier routes to help for children and young people*. London: Office of the Children's Commissioner.

Cotmore, R. & Fernandes, P. (2008) Review of literature on evaluating helpline outcomes. Unpublished report. London: NSPCC.

Coulter, A. (2006) *Engaging patients in their healthcare: How is the UK doing relative to other countries?* Oxford: Picker Institute Europe.

Coveney, C.M., Pollock, K., Armstrong, S. & Moore, J. (2012) Callers' experiences of contacting a national suicide prevention helpline: Report of an online survey. *Crisis: The Journal of Crisis Intervention and Suicide Prevention,* 33, pp.313–324.

Cronin, C., Sood, S. & Thomas, D. (2017) From innovation to transcreation: Adapting digital technologies to address violence against children. *Child Abuse Review,* 26, pp.215–229.

Cullen, M.A., Simpson, C., Bubley, R., Pike, M., Cullen, S. & Allen, B. (2017) *Delivering family interventions through social investment? Sharing the learning from a voluntary sector case study*. Addaction.

Daly, M., Bray, R., Bruckauf, Z. et al. (2015) *Family and parenting support: Policy and provision in a global context*. Florence: UNICEF Office of Research – Innocenti.

Danby, S. & Emmison, M. (2012) Kids, counsellors and troubles-telling: Morality-in-action in talk on an Australian children's helpline. *Journal of Applied Linguistics and Professional Practice,* 9, pp.263–285.

DfE (Department for Education) (2017a) *Characteristics of children in need: 2016 to 2017*. London: DfE.

DfE (2017b) *Children looked after in England*. London: DfE.

Dinh, T., Farrugia, L., O'Neill, B., Vandoninck, S. & Velicu, A. (2016) *Insafe helplines: Operations, effectiveness and emerging issues for internet safety helplines*. European Schoolnet (EUN), Insafe and the Kaspersky Helpline Fund.

Donaldson, R. & Irwin, J. (2017) *YoungMinds: Parents Helpline evaluation report*. Fiveways.

Dottridge, M. (2008) *Kids abroad: Ignore them, abuse them or protect them?*. Geneva: Terre des Hommes.

Downing, J. & Cook, P.A. (2006) *Evaluation of young people's contraceptive and sexual health services in Knowsley*. Liverpool: Public Health Institute, Liverpool John Moores University.

Drolet, A.L. & Morrison, D.G. (2001) Do we really need multiple-item measures in service research? *Journal of Service Research,* 3, pp.196–204.

EPPI-Centre (2018) *Definitions*. http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=334

Fernandes, P. (2006) *Evaluating children's participation in decision making: A case study of a child helpline in India*. London: Department of Social Policy, London School of Economics and Political Science.

Finn, J. & Hughes, P. (2008) Evaluation of the RAINN National Sexual Assault Online Hotline. *Journal of Technology in Human Services,* 26, pp.203–222.

Fry, D., Casey, T. & NSPCC (2017) *Bringing the global to the local: Review of global trends in the prevalence and services for child maltreatment in order to inform research, policy and practice in England*. London: Early Intervention Foundation.

Fukkink, R.G. & Hermanns, J.M.A. (2009) Children's experiences with chat support and telephone support. *Journal of Child Psychology and Psychiatry,* 50, pp.759–766.

Gallagher, M. (2013) *Helpline highlight: Northern Ireland: A year in review: April 2011–March 2012*. London: NSPCC.

Gallagher, M. (2014) *Speaking up for Scotland's children: The NSPCC's helpline in Scotland: A year in review April 2012–March 2013*. Edinburgh: NSPCC.

Goldsmith, C. (2017) *At the end of the line: Independent evaluation of the Pact Helpline*. North Research, Training and Development.

Gough, D., Thomas, J. & Oliver, S. (2012) Clarifying differences between review designs and methods. *Systematic Reviews,* 1, pp.1–9.

Grant, M. & Booth, A. (2009) A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal* 26, 91–108.

Grove, C., Reupert, A. & Maybery, D. (2015) Peer connections as an intervention with children of families where a parent has a mental illness: Moving towards an understanding of the processes of change. *Children and Youth Services Review,* 48, pp.177–185.

Haner, D. & Pepler, D. (2017) Adolescents show positive changes in distress and hope after single session, post-based, anonymous counselling at Kids Help Phone. *Children and Youth Services Review,* 82, pp.207–213.

Hanley, T. (2009) The working alliance in online therapy with young people: Preliminary findings. *British Journal of Guidance & Counselling,* 37, 3, pp.257–269.

Hepburn, A. (2005) 'You're not takin' me seriously': Ethics and asymmetry in calls to a child protection helpline. *Journal of Constructivist Psychology,* 18, 3, pp.253–274.

Hudson, B. (2005) 'Not a cigarette paper between us': Integrated inspection of children's services in England. *Social Policy and Administration,* 39, pp.513–527.

Hutchinson, D. (2012) *Helpline highlight: A year in review*. London: NSPCC.

Hutson, N. & Cowie, H. (2007) Setting up an email peer support scheme. *Pastoral Care in Education,* 25, 4, pp.12–16.

Karver, M.S., Tarquini, S.J. & Caporino, N.E. (2010) The judgment of future suicide-related behavior: Helpline counselors' accuracy and agreement. *Crisis: The Journal of Crisis Intervention and Suicide Prevention,* 31, pp.272–280.

King, R., Bambling, M., Reid, W. & Thomas, I. (2006) Telephone and online counselling for young people: a naturalistic comparison of session outcome, session and impact therapeutic alliance. *Counselling & Psychotherapy Research,* 6, 3, pp.175–181.

Kitchingman, T.A., Wilson, C.J., Caputi, P., Woodward, A. & Hunt, T. (2015) Development and evaluation of the Telephone Crisis Support Skills Scale. *Crisis: The Journal of Crisis Intervention and Suicide Prevention,* 36, pp.407–415.

Law, J., Haner, D., Simon, A. et al. (2015) *Proof positive: Evaluation of Kids Help Phone's phone and live chat counselling services*. Kids Help Phone.

Mental Helplines Partnership (2003) *Help on the line. What commissioners and funders need to know about mental health helplines*. London: Telephone Helplines Association.

Ministry for Women & Netsafe (2017) *Insights into digital harm: The online lives of New Zealand girls and boys*. Wellington, New Zealand: Ministry for Women and Netsafe. Available at: http://women.govt.nz/documents/insights-digital-harm-online-lives-new-zealand-girls-and-boys

Mishara, B.L., Chagnon, F., Daigle, M. et al. (2007) Which helper behaviors and intervention styles are related to better short-term outcomes in telephone crisis intervention? Results from a silent monitoring study of calls to the US1-800-SUICIDE Network. *Suicide and Life-Threatening Behavior,* 37, pp.308–321.

Mokkenstorm, J.K., Eikelenboom, M., Huisman, A. et al. (2017) Evaluation of the 113Online suicide prevention crisis chat service: Outcomes, helper behaviors and comparison to telephone hotlines. *Suicide and Life-Threatening Behavior,* 47, pp.282–296.

Morgan, K., Chakkalackal, L. & Cyhlarova, E. (2012) *Life lines: Evaluation of mental health helplines*. London: Mental Health Foundation.

Morton, S. (2015) Progressing research impact assessment. *Research Evaluation,* 24, pp.405–419.

Nicholas, L. & Broadbent, S. (2015) *Collective intelligence in patient organisations*. London: National Endowment for Science, Technology and the Arts (NESTA).

NSPCC (2008) *A pocket history of the NSPCC*. London: NSPCC.

NSPCC (2016) *Childline: 30 years of listening to children*. London: NSPCC.

Ofsted (2004) *Every Child Matters: Inspection of children's services; Key judgements and evidence*. London: Ofsted.

Phillips-Howard, P., Madden, H., Mason, L. et al. (2010) *The WiSHing Project: Wirral Sexual Health, health and wellbeing needs assessment of young people*. Liverpool: Public Health Institute, Liverpool John Moores University.

Potter, J. & Hepburn, A. (2005) Discursive psychology as a qualitative approach for analysing interaction in medical settings. *Medical Education,* 39, pp.338–344.

Potter, J. & Hepburn, A. (2012) Somewhere between evil and normal: Traces of morality in a child-protection helpline. *Journal of Applied Linguistics and Professional Practice,* 9, pp.245–262.

Rantzen, E. & NSPCC (2017) *Not alone anymore: Childline annual review 2016/17*. London: NSPCC.

Reubsaet, A., Lechner, L. & de Vries, H. (2006) The Dutch cancer information helpline: More critical patients after 10 years. *Patient Education & Counseling,* 63, pp.215–222.

Rhee, W.K., Merbaum, M., Strube, M.J. & Self, S.M. (2005) Efficacy of brief telephone psychotherapy with callers to a suicide hotline. *Suicide & Life-threatening Behavior,* 35, 3, pp.317–328.

Rogers, P.J. (2008) Using programme theory to evaluate complicated and complex aspects of intervention. *Evaluation,* 14, pp.29–48.

Santor, A., Poulin, C., Leblanc, J. & Kusumakar, V. (2007) Online health promotion, early identification of difficulties, and help seeking in young people. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46, pp.50–59.

Scottish Government (2008) *Child protection line campaign evaluation: Report for the Scottish Government and the Union*. Edinburgh: TNS System Three.

Sharratt, K., Porter, J. & Truman, C. (2014) 'Nowhere else to turn' – Key findings from an evaluation of the offenders' families helpline. *Prison Service Journal*, pp.30–35.

Shelter (2011) *Keys to the future: An independent evaluation of Shelter Children's Services, England and Scotland*, Shelter.

Sindahl, T.N. (2013) *Chat counselling for children and youth – A handbook*. Child Helpline International.

Starks, L., Kendall, S. & Whitley, J. (2012) *Reach: A new model of intervention for children before, during and after they run away: Executive summary*. Sandbach: Railway Children.

Statham, J. & Carlisle, J. (2004) *Calling for help: The role of telephone helplines in child protection. A study for the Children's Safeguards Unit*. London: Thomas Coram Research Unit.

Tyson, P., Law, C., Reed, S., Johnsey, E., Aruna, O. & Hall, S. (2016) Preventing suicide and self-harm. *Crisis: The Journal of Crisis Intervention & Suicide Prevention,* 37, pp.353–360.

UNICEF (2007) *Implementation handbook for the Convention on the Rights of the Child*. Geneva: UNICEF Regional Office for Europe.

University of Central Lancashire and NCB (National Children's Bureau) Research Centre (2016) *Independent advocacy: Impact and outcomes for children and young people*. Office of the Children's Commissioner.

University of Edinburgh, Zimbabwe C and Zimbabwe UNCsFU (2016) *A secondary analysis of data from Childline Zimbabwe*. Understanding Violence Against Children in Zimbabwe Data Series, Harare: UNICEF.

van Dolen, W. & Sindahl, T.N. (2017a) *Impact assessment: A guide for child helplines*. Amsterdam: Child Helpline International

van Dolen, W. & Sindahl, T.N. (2017b) *Impact report: Child helplines*. Amsterdam: Child Helpline International.

van Dolen, W. & Weinberg, C. (2017) Child helplines: How social support and controllability influence service quality and well-being. *Journal of Services Marketing,* 31, pp.385–396.

Vogel, I. (2012) *Review of the use of 'Theory of Change' in international development*. London: Department for International Development.

Wilson, T. & Haner, D. (2012) *Live chat counselling pilot*. Kids Help Phone.

Wood, M., Wanless, P. & NSPCC (2016) *Safer childhoods, brighter futures: Keeping children safe today and for generations to come: Annual report and accounts 2015/*16. London: NSPCC.