Childonomics – Methodology for Appraising the Return on Investment of Social Services for Children and Families

Adrian Gheorghe, Joanna Rogers, Jean Anne Kennedy and Volodymyr Kuzminskyi

December 2017
Acknowledgements

Many thanks to Andy Bilson, Richard Carter, Jennifer Davidson, Chris Desmond, Andrew Goudie, Jana Hainsworth, Maria Herczog, Chris Rayment, Nicholas Rees, and Purti Sharma, who all provided comments on earlier drafts of this paper. Thank you also to other members of the Childonomics Advisory Group, who have made valuable contributions. We are also grateful to Clara Picanyol from Oxford Policy Management, who peer-reviewed this final version of the paper.
Executive summary

The Childonomics research project has developed a methodology that can help decision makers, child care specialists, social workers, researchers, child rights advocates, and other interested parties reflect on the long-term social and economic return of investing in children and families within a given national or sub-national context. Childonomics is trying to do something very challenging and for which there are no standard models. The methodology paper presented here should be read in conjunction with the overarching conceptual framework (Childonomics: A Conceptual Framework; Bilson et al., 2017)¹ and the report presenting the results of testing the methodology in Romania and Malta (Childonomics: Malta and Romania, lessons learned from applying the Childonomics conceptual framework and methodology in practice; Rogers et al., 2017).

This working paper describes a methodology for assessing the value of different types of service by enabling consideration of the costs of different services and approaches to supporting children and families in vulnerable situations – or at risk of entering into vulnerable situations – and the kinds of outcomes that interaction with those services might be expected to achieve for the child and family, as well as for the community and society as a whole. The methodology focuses on the impact dimension in the Childonomics conceptual framework, with an emphasis on valuing the investments and outcomes associated with different services. Outcomes go beyond the perspective of the child to include parents and the extended family, communities, and society at large. The project and methodology adopt a rights-based foundation and outcomes focus as a basis for understanding the investment in children and families that is required to ensure the well-being of children, with a particular emphasis on supporting children, families, and communities in order to prevent and reduce any form of developmental delay, harm and, especially, the unnecessary separation of children from their parents or families.

The methodology proposes an approach to economic modelling that can be used in a number of ways to inform decision making, advocacy, research and practice development. A full evaluation methodology, which allows for the simultaneous comparison of costs and outcomes across several alternatives, is most in line with the scope of Childonomics and the methodology most closely aligned with its aims is cost-consequence analysis (CCA). CCA presents costs and outcomes side by side in a disaggregated manner; it is a form of cost-effectiveness analysis which presents the range of benefits identified alongside costs incurred without aggregating them in a single metric (e.g. a cost-effectiveness ratio), leaving the users of the methodology to incorporate their own considerations in judging the merits of the intervention or programme.

The logic of the methodology is summarised in Figure 1 and outlined narratively thereafter.

**Figure 1: The logic of the Childonomics methodology**

<table>
<thead>
<tr>
<th>Steps</th>
<th>Key activities</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conduct analysis</td>
<td>Stakeholder consultation, Policy analysis, Literature review</td>
<td>Outcome matrix</td>
</tr>
<tr>
<td>Assess context</td>
<td>Stakeholder consultation, Literature review(s), Primary data collection, Economic modelling, Other modelling forms</td>
<td>Results matrix</td>
</tr>
<tr>
<td>Establish scope</td>
<td>Validation workshop</td>
<td>Summative narrative</td>
</tr>
<tr>
<td>Specify and measure outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specify and measure costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimate net costs and outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assess strength of evidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construct narrative</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Step 0: Assess context**

This entails being able to answer the following questions:

- What are the policies, strategies, legislation, and rules and regulations governing the (areas of interest in the) child and family services sector?
- Who are the main stakeholders?
- What are the resource allocation mechanisms in the sector?
- What are the main routine data sources?

**Step 1: Establish scope**

This entails making decisions and being able to give clear answers to the following questions:

- What is the main purpose of the analysis and who is the main audience?
- Which services (or types of service) are being compared?
- What is the economic perspective of the analysis?
- What is the time horizon for considering costs and effects?

**Step 2: Specify outcomes at each level of the Childonomics framework**

Outcomes are specified at four levels: child, parents/family, community and society. In order to identify the outcomes and their indicators, there are two perspectives to consider:

- **A data-informed perspective**, which involves using available datasets, consulting official and (routinely collected) statistics, conducting literature reviews, etc.
**A beneficiary-informed perspective**, which entails engaging service beneficiaries (e.g. children and their families) in a structured dialogue that reveals what matters to them. In a broader sense, it is essential to engage all relevant stakeholders in this process – an element that the Social Return on Investment (SROI) methodology also features strongly.

Reviewing available theory and data, as well as consulting service beneficiaries, will result in the completion of an **outcomes matrix** (an illustrative example of such a matrix is shown in Table 1).

### Table 1: Illustrative outcomes matrix from Romania pilot (day centres for children with disabilities)

<table>
<thead>
<tr>
<th>Domain</th>
<th>Indicator</th>
<th>Data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outputs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outputs for service users</td>
<td>Child development</td>
<td>Development, speech/communication assessments; case files; interviews with service providers and parents</td>
</tr>
<tr>
<td></td>
<td>Increased capacity of parents to provide care</td>
<td>Interviews with carers and parents; interviews with children and young people; case files</td>
</tr>
<tr>
<td>Outputs for service users</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child</td>
<td>Children prepared for school/education</td>
<td>Official statistics, service data and case files; interviews with young people (and former service users); education records</td>
</tr>
<tr>
<td>Family</td>
<td>Parents employed and relationships stable</td>
<td>Service data and case files</td>
</tr>
<tr>
<td>Outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community</td>
<td>Reduced separation of children from parents</td>
<td>Official statistics, service data and case files</td>
</tr>
<tr>
<td>Society</td>
<td>Increased education and inclusion of children with disabilities</td>
<td>Official statistics, service data and case files</td>
</tr>
</tbody>
</table>

Source: Authors, based on document review and interviews with staff and service users at two day centres for children with disabilities in Romania, 2017.

**Step 3: Measure outcomes**

Once the outcomes are specified, indicator values will be determined through various types of data and modelling. In this context, **modelling** can be used to describe the effect of each intervention (the comparator and alternatives) on each outcome. Modelling is to be understood here broadly as a **process** comprising several generic steps:

---

^2 AROPE (‘at risk of poverty and exclusion rate’) is an indicator used by the European Union in monitoring the implementation of 2020 strategies in EU member states.
• **Identifying the relevant data.** Obtaining quantitative or qualitative information on the target outcome indicators. This could be data that are already available (for example in databases, published studies, or reviews of the literature) or data collected as part of the analysis.

• **Assessing the extent to which data constitutes evidence for the said outcomes.** This implies making a judgement on validity and generalisability. For validity, this means determining the extent to which observed effects are due to exposure to the intervention. For generalisability, the focus is on determining the alignment between how and where the data were obtained and the scope of the present analysis.

• Potentially **processing the data**, e.g. by combining with other data, to obtain new insights. For quantitative data, this is often done by specifying mathematical relationships. For qualitative data, triangulation is a well-used social science approach that brings together several types of data or sources to create a richer, more rounded understanding of the issue at hand.

**Step 4: Calculate costs**

Three 'core' types of costs will be calculated:

• **Total costs** of delivering the service/implementing the policy scenario.

• **Total costs per beneficiary** (child, family, etc.).

• **Marginal cost per beneficiary**, i.e. the cost of delivering the service to one additional beneficiary (child, family, etc.) once the overhead costs have been accounted for.

These can be complemented with other quantities, depending on the specificities of the analysis.

**Step 5: Calculate and present net costs and outcomes**

Net costs and outcomes are determined as **differences for each indicator between alternative(s) and the comparator**. This operation is straightforward for quantitative indicators. However, for some outcomes, it may be the case that only a qualitative judgement can be made due to lack of appropriate means of measurement, e.g. on parent knowledge and competency or progress in social skills. An additional consideration is the **strength of evidence** supporting the net impacts. We propose here a simple taxonomy of the strength of evidence for each net impact (cost or outcome) based on the two considerations in Step 3, namely validity and generalisability. Evidence for impact can be classified as **strong, moderate or indicative** based on a combination of scores along the two dimensions.

Net costs and outcomes are presented in a **results matrix** (see Table 2), together with information on the strength of evidence supporting the net impacts. The matrix is meant (1) to offer analysts and stakeholders a comprehensive view of all the findings, and (2) to guide them towards applying their own judgement on the comparison by offering information on net impact and the strength of supporting evidence.
### Table 2: Example results matrix

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Domain</th>
<th>Indicator (examples)</th>
<th>Comparator – Family support services</th>
<th>Alternative 1 – Residential care</th>
<th>Alternative 2 – Foster care</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Impact</td>
<td>Impact</td>
<td>Difference from comparator</td>
</tr>
<tr>
<td>Outputs</td>
<td>Outputs</td>
<td>Contact time per beneficiary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Child-level</td>
<td>Psychological well-being</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Family</td>
<td>Number of prevented separations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Community</td>
<td>% variation in homelessness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Society</td>
<td>Productivity in adulthood</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Risk of poverty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Costs</td>
<td>Total cost</td>
<td>Total cost of programme per year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total cost per beneficiary</td>
<td>Total cost per child</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total cost per family</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marginal cost per beneficiary</td>
<td>Marginal cost per child</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Marginal cost per family</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© Oxford Policy Management
Step 6: Validate findings with stakeholders and create a summative narrative

In the final step, a validation workshop is held where stakeholders who have been involved in the process from its early stages engage with the results matrix. The workshop aims to construct a shared understanding of the results among those involved and to use it as a foundation for results interpretation. This would follow a series of steps:

1. **Reviewing, understanding and validating the findings**: this is conducted by the team responsible for conducting the analysis using the five steps outlined above, who present the analytical process and the results matrix.

2. **Prioritising and validating the impacts**: this takes the form of a guided discussion among the workshop participants on priorities and the strength of evidence, as well as a reflection on the analysis of differences between the comparators.

3. **Reflecting on the results matrix and creating a summative narrative**: this is a final group discussion, arriving at a joint conclusion on the relative value of the examined interventions and factors impacting their relative value based on the available evidence, stakeholder prioritisation and interpretation of contextual issues identified by stakeholders.

The final output is the results matrix with a summative narrative that presents in a distilled form the analysis, the discussion from the validation workshop, and the final joint stakeholder and research team conclusion on the relative value of the examined interventions. The summative narrative provides an explanation of how the net outcomes for each alternative are valued by stakeholders and sets out any contextual issues that may affect the interpretation of the results matrix.

**Conclusions**

The main challenges in implementing the methodology are primarily in defining outcomes while balancing value-laden questions of the relative worth of outcomes for different stakeholders. The methodology is dependent on the quality and validity of data that is available. The strengths of the methodology are its ability to engage stakeholders across institutional and organisational boundaries in a joint examination of assumptions, costs, and outcomes, while giving due consideration to the strength of the available evidence to support the conclusions reached.

Overall, the methodology aims to foster evaluative thinking in the field of child services through structured, system-wide, robust, and context-specific analyses that balance a rights-based foundation and outcomes focus with fundamental elements of evidence-informed decision making in the public sector.
# Table of contents

Acknowledgements ................................................. i
Executive summary .............................................. ii
List of tables, figures and boxes ................................. ix
List of abbreviations ............................................ x

## 1 Introduction ................................................. 1

## 2 Background to measuring the value of social services  4

### 2.1 Costs .................................................. 4

### 2.2 Outcomes ............................................. 6

### 2.3 Combining costs and outcomes .......................... 11

## 3 Proposed methodology ..................................... 16

### 3.1 The *Childonomics* approach ........................... 17

### 3.2 Analysis .............................................. 20

### 3.3 Strength of evidence .................................. 26

### 3.4 The summative narrative: what do the findings mean and to whom? 29

## 4 Final considerations ....................................... 33

## Annex A Examples of approaches to grading evidence  34

### A.1 Oxford Centre for Evidence-Based Medicine (OCEBM) 34

### A.2 California Evidence-Based Clearinghouse for Child Welfare 36

### A.3 Petticrew and Roberts (2003) .......................... 37

### A.4 Standards of Useful Evidence, Dartington Social Research Unit 38
List of tables, figures and boxes

Table 1: Illustrative outcomes matrix from Romania pilot (day centres for children with disabilities) ................................................................. iv
Table 2: Example results matrix ........................................................................................................................................................ vi
Table 3: OECD example – selection of child well-being indicators ................................................................. 10
Table 4: Malta pilot example of outputs and outcomes ........................................................................................................... 11
Table 5: Types of economic evaluation methodology ................................................................................................. 12
Table 6: Illustrative outcome matrix from Romania pilot (day centres for children with disabilities) ......................................................................................................... 23
Table 7: Summary of strength of evidence categories ................................................................................................. 28
Table 8: Definitions of strength of evidence categories ................................................................................................. 29
Table 9: Illustrative results matrix .................................................................................................................................................. 30

Figure 1: The logic of the Childonomics methodology ........................................................................................................ iii
Figure 2: Childonomics conceptual framework illustrated with examples drawn from the Guidelines on Alternative Care for Children and global literature review ........................................................................ 2
Figure 3: Example of categories of outcomes for children receiving highly specialised services ... 7
Figure 4: Example of an outcomes monitoring system ............................................................................................................. 8
Figure 5: Example of visualisation of progress in achieving outcomes ...................................................................... 8
Figure 6: VfM 3E framework applied to social protection systems ......................................................................................... 14
Figure 7: Methodology outline .................................................................................................................................................. 19

Box 1: The Gates Reference Case ................................................................................................................................................ 16
Box 2: Incorporating effect size .................................................................................................................................................. 24
Box 3: Summary of approaches and data sources for outcome specification and estimation used in Malta and Romania .................................................................................................................................................. 25
Box 4: Final output from the Childonomics CCA .................................................................................................................. 32
## List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AROPE</td>
<td>At risk of poverty and exclusion rate</td>
</tr>
<tr>
<td>CBA</td>
<td>Cost–benefit analysis</td>
</tr>
<tr>
<td>CCA</td>
<td>Cost-consequence analysis</td>
</tr>
<tr>
<td>CEA</td>
<td>Cost-effectiveness analysis</td>
</tr>
<tr>
<td>CEE</td>
<td>Central and Eastern European Countries</td>
</tr>
<tr>
<td>CIS</td>
<td>Commonwealth of Independent States</td>
</tr>
<tr>
<td>OCEBM</td>
<td>Oxford Centre for Evidence-Based Medicine</td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Co-operation and Development</td>
</tr>
<tr>
<td>OPM</td>
<td>Oxford Policy Management</td>
</tr>
<tr>
<td>PSSRU</td>
<td>Personal and Social Services Research Unit</td>
</tr>
<tr>
<td>QALY</td>
<td>Quality adjusted life year</td>
</tr>
<tr>
<td>ROI</td>
<td>Return on Investment</td>
</tr>
<tr>
<td>SROI</td>
<td>Social Return on Investment</td>
</tr>
<tr>
<td>UNCRC</td>
<td>United Nations Convention on the Rights of the Child</td>
</tr>
<tr>
<td>VfM</td>
<td>Value for Money</td>
</tr>
</tbody>
</table>
1 Introduction

The *Childonomics* research project has developed a methodology that can help decision makers, specialists, researchers, child rights advocates, and other interested parties reflect on the long-term social and economic return of investing in children and families within a given national or sub-national context. The methodology enables consideration of the costs of different services and approaches to supporting children and families in vulnerable situations – or at risk of entering into vulnerable situations – and links them to the kinds of outcomes which interaction with those services might be expected to achieve for the child and family, as well as for the wider community and society as a whole. The project and methodology use a rights-based foundation and an outcomes focus as a basis for understanding the investment in children and families that is required to ensure the well-being of children, with a particular emphasis on supporting children, families and communities in order to prevent and reduce any form of developmental delay, harm, and especially the unnecessary separation of children from their parents or families. This methodology should be read together with the *Childonomics* conceptual narrative and the report presenting the results of testing the methodology in Romania and Malta as part of the *Childonomics* project in 2016–2017.4

In the first phase of *Childonomics*, a conceptual framework was developed and presented to stakeholders to illustrate the conceptual narrative set out in *Childonomics: A Conceptual Narrative* (Bilson et al., 2017). The conceptual framework (Figure 2) attempts to illustrate, based on the Guidelines on Alternative Care for Children5 and a review of the global evidence base, the range of services that children and families can access which may influence outcomes (on the right side of the diagram) at different levels for:

- The child (e.g. improved development, education outcomes or employment opportunities);
- The parents or family (improved ability to provide care, ensure child development, provide a nurturing environment);
- The community (e.g. less crime committed by young people, improved school attendance); and
- Society at large (e.g. greater engagement of young people in employment, education or training; lower rates of suicide among young people; lower rates of unplanned or juvenile pregnancy; fewer children in out-of-home care).

---

Figure 2: *Childonomics* conceptual framework illustrated with examples drawn from the Guidelines on Alternative Care for Children and global literature review

The conceptual framework offers a way of mapping services and programmes in any given national or sub-national setting, including universal services that serve all children and families – such as education and health services – or social assistance services that are targeted to particular populations, e.g. low-income households with children. Specialised services are even more targeted to meeting specific needs, e.g. services for children with disabilities, and higher intensity, more specialised services for children and families facing more complex challenges,
such as the risk of physical or sexual abuse. The bold line towards the bottom of the diagram indicates the point where children leave the care of their families and enter formal care.  

The framework assumes that a range of investments can have an impact on children and the ability of their families to provide care for them. All children, regardless of the setting where they live, should have access to education, health, and other universal services. Children and families facing challenges (e.g. poverty, disability, health problems, etc.) may require additional support from targeted services (social assistance, disability budgets, employment programmes). Children and families experiencing extreme or complex challenges may require specialised services (parent education programmes, social accompaniment services) or highly specialised services aimed at preventing or addressing the impact of abuse, violence, or neglect.

If children are outside the care of their parents or family due to the death of parents or because parents or carers cannot look after them adequately even with support from specialised services, then they require alternative care services (kinship care, foster care, small group homes, other types of residential care) that are suitable for meeting their needs and can support positive outcomes. The conceptual framework offers a basis for comparing the outcomes resulting from these different alternative care options; it also offers a way of comparing investment in alternative care with investments in services when children are still in the care of their families. The framework offers a way of considering the outcomes that might result from adjusting the level of investment in different parts of the system, for example increasing expenditure in one part of the system (e.g. targeted or specialised services to support children and families living in poverty) to the levels found in another part of the system (e.g. alternative care services).

The idea of the conceptual framework is to examine the investment being made and to link it as far as possible to the outcome that the investment can help to provide. The model takes a broad view of the types of services that can be the object of inquiry and approaches them from several perspectives, namely:

- **availability** – the extent to which various types of services exist in the given setting;
- **accessibility** – the extent to which services that exist are used by the target population; and
- **impact** – the extent to which various degrees of investment in such services generate impact.

This working paper describes a methodology for assessing the value of different types of service and focuses particularly on the impact dimension in the *Childonomics* conceptual framework, specifically on valuing the investments and outcomes associated with different services. In line with the conceptual framework, outcomes under consideration go beyond the perspective of the individual child to include parents and the extended family, communities and society at large. The methodology provides an approach to economic modelling that can be used in a number of ways to inform decision making, advocacy or practice development. The methodology was piloted and refined based on analytical work and field visits (February–October 2017) in Romania and Malta and through discussions and review by members of the *Childonomics* Advisory group.

This paper continues with an overview of some considerations relevant to measuring the value of social services (Section 2) and an outline of the proposed methodology (Section 3).

---

6 As defined in Article 29.b) ii) of the UN Guidelines on Alternative Care for Children (2009): ‘Formal care: all care provided in a family environment which has been ordered by a competent administrative body or judicial authority, and all care provided in a residential environment, including in private facilities, whether or not as a result of administrative or judicial measures.’
2 Background to measuring the value of social services

There are three fundamental elements to consider when assessing the value of a generic service or intervention from an economic perspective:

- **Costs**: this refers primarily to the resources required to deliver the service, but may also include expenditure incurred by other stakeholders (e.g. transport costs incurred by beneficiaries when accessing the service, such as transport costs), depending on the perspective taken in the analysis. The main activities related to cost estimation are identifying, measuring, valuing, and comparing the value of resource items associated with delivering and using the service.

- **Outcomes**: this refers to the eventual benefits to society, community and/or individual families and children that the proposed social services are intended to achieve. However, negative and unintended results can also occur and they need to be accounted for in the analysis. As with the costs, the main activities are identifying, measuring, valuing, and comparing outcomes.

- **Value**: here this means using information on costs and outcomes to inform a judgement on the value of the service in question.

An in-depth methodological discussion on these elements, such as cost estimation, is beyond the scope of this paper and there is ample literature available on the topic, which is referenced in subsequent sections as needed. In what follows, we discuss for each element several issues that are pertinent to Childonomics.

2.1 Costs

Several classifications of costs are available, taking into account various accounting or economic considerations such as cost behaviour (fixed cost vs marginal cost), cost occurrence (capital cost vs recurrent cost), or traceability to ‘cost objects’ (direct costs vs indirect costs). Choosing an appropriate costing taxonomy depends primarily on the decision problem at hand.

We use here a generic, simple classification to illustrate the main types of cost relevant to social services:

- **Costs incurred by the service provider**:
  - Direct service costs, i.e. the costs of delivering the service or the intervention from the perspective of the provider of the service.

- **Costs not incurred by the service provider**:
  - Beneficiary out-of-pocket expenditures directly related to the service (e.g. fees for a foster child’s consultation with a specialist not covered by the service provider) or not (e.g. transport to the service provider);
  - Costs incurred by other public sector branches;
  - Cost of informal care, including imputed loss of carers’ earnings;
  - Productivity costs;
  - Future service costs; and
  - Transfer payments, e.g. cash transfers to vulnerable individuals.
The list above is not exhaustive; additionally, not all cost categories are relevant to all costing problems. The costing perspective and data availability/reliability will determine which costs to account for and which to leave out. For example, for an analysis conducted from the perspective of the service provider, estimating direct service-related costs is sufficient; for an analysis conducted from a broad societal perspective, productivity costs and the cost of informal care must be included because they have broader societal implications.

Estimating direct service costs usually requires the most intensive analytical effort. Two broad estimation approaches can be used:

- **A top-down approach**, which entails assembling the relevant expenditure and dividing it by the unit of activity or output. This approach is relatively simple to apply, particularly for existing interventions, as it relies on routine expenditure data and is expedient when the average cost per unit is required.

- **A bottom-up approach**, which entails identifying, measuring, and valuing the individual resource items that make up a service and then dividing their total value by the unit of activity or output. This is a more effort-intensive approach and also leads to more precise estimates. It is particularly suitable for ‘new’ interventions which have not yet been institutionalised in the given setting, and allows replication of the cost estimation elsewhere, e.g. by varying the quantities or valuation of individual inputs.

Irrespective of the approach, the following types of cost make up the provider-incurred costs of any generic service:

- One-off costs, such as: capital costs (building, land, equipment, furniture) or one-time set-up costs (initial recruitment and training of staff, etc.); and

- Recurrent costs: human resources, consumables, administration and management (e.g. cleaning, utilities, maintenance, rents, communication, security, logistics, accounting, etc.).

A tutorial on applying these approaches and classifications in practice is beyond the scope of this document. A good amount of methodological guidance is available on the topic, including specifically for costing child and family social services; strongly recommended reading is the Personal and Social Services Research Unit (PSSRU) guide. We highlight instead two considerations particularly relevant for estimating and interpreting the costs of child services:

- Introducing a new service may actually increase the number of users or beneficiaries. As such, estimating the future total costs of a service based on current or historical demand for a similar service may lead to underestimates. This potential increase in beneficiaries needs to be reflected in the estimations of future costs.

- Increasing costs over time, particularly in the early stages of introducing a service, may be difficult to interpret. For example, they may signal capacity to benefit, positive outcomes and long-term savings, as Boddy et al. (2012) found in a study of families accessing intensive family support whose health needs were largely unmet prior to accessing the service. As such, analysing costs in isolation from outcomes can be misleading.

---


2.2 Outcomes

The generic challenge with outcomes in social services relates to difficulties in measuring them directly because suitable indicators either cannot be specified or relevant data are unavailable to inform them. Such measurement challenges are equally relevant to individual-level (child) and broader (family, community, and societal) outcomes.

At the individual level, for example, case management is an approach to coordinating the provision of specialised, highly specialised, and alternative social services based on an assessment of need. By aligning the process of needs assessment with the setting up of relevant targets and objectives for interventions, it is possible to develop indicators that can identify progressive levels of meeting needs or of clients' abilities to achieve specific tasks. Thus, needs assessment forms a fundamental basis, not only for service planning and service provision, but also for the monitoring and evaluation of outcomes for each client.

Unfortunately, such outcome indicators are not easy to determine because of difficulties with defining 'outcomes' and setting up a framework that links different outcome categories with the needs assessment instrument. There are few frameworks that have found widespread support and become established as frameworks for assessing outcomes for children. Here we provide an overview of one or two examples from the literature to illustrate the challenges of creating such a framework rather than to provide a recommended model or guide.

One approach is to conceptualise outcomes as a profile of categories, such as the twelve categories for children in need proposed in the UK Common Assessment Framework. As illustrated in Figure 3, all categories of outcome are vectors (they have magnitude and direction) and are arrayed in parallel.

Outcome profiles can be synthesised in an outcomes monitoring system, as shown in Figure 4. For convenience, all categories of outcomes are marked with a value ranging from 0 to 100. The model allows removing, inputting, and replacing one category of outcomes with another where necessary, making it relatively easy to adapt to existing assessment forms and particular social services. It has further advantages: it is visual and easy to interpret; it provides a basis for monitoring change in outcomes over time, for example using parallel graphs (Figure 5); and it can be created using a standard Excel spreadsheet.

10 See www.education.gov.uk/childrenandyoungpeople/strategy/integratedworking
### Figure 3: Example of categories of outcomes for children receiving highly specialised services

<table>
<thead>
<tr>
<th>Category</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. General health</td>
<td>No diseases - Chronic illness</td>
</tr>
<tr>
<td>2. Physical development</td>
<td>Normal - Poor</td>
</tr>
<tr>
<td>3. Speech, language and communication</td>
<td>Functional - Disabled</td>
</tr>
<tr>
<td>4. Emotional and social development</td>
<td>Satisfied - Dissatisfied</td>
</tr>
<tr>
<td>5. Behavioural development</td>
<td>Common - Disorder</td>
</tr>
<tr>
<td>6. Identity, self-esteem, self-image</td>
<td>Normal, positive - Abnormal, negative</td>
</tr>
<tr>
<td>7. Family and social relationships</td>
<td>Strong, high level of attachment - Weak, lack of attachment</td>
</tr>
<tr>
<td>8. Self-care skills and independence</td>
<td>Confident - Fearful</td>
</tr>
<tr>
<td>9. Understanding, reasoning and problem-solving</td>
<td>High-level - Low-level</td>
</tr>
<tr>
<td>10. Participation in learning, education</td>
<td>Active - Passive</td>
</tr>
<tr>
<td>11. Progress and achievement in learning</td>
<td>Achieving - Not achieving</td>
</tr>
<tr>
<td>12. Aspirations</td>
<td>Ambitions - Apathies</td>
</tr>
</tbody>
</table>

Source: Authors, based on Starfield\(^{11}\).

---

Figure 4: Example of an outcomes monitoring system

Source: P4EC CEE/CIS Consultancy Group, 2013.

Figure 5: Examples of visualisation of progress in achieving outcomes
Going beyond the individual level, an approach to assess the broader outcomes of social services uses the concept of social impact. Social impacts are the effects of services on society or the community and the well-being of individuals and families. Such impacts can be measured in terms, for example, of changes in levels of health, crime, sustainability, education, community cohesion, and diversity. Analytically, net impact is calculated by subtracting what would have happened anyway from measured outcomes.

The challenge is to find robust indicators for measuring outcomes. There is no unique, universally accepted range of outcomes and indicators for measuring these outcomes that emerges from the academic literature. Some authors and international organisations such as the Organisation for Economic Co-operation and Development (OECD) suggest using the concept of ‘child well-being’ for measuring the results of child-focused policies across countries. For measuring child well-being it is suggested to use a number of indicators, which are grouped into six domains: material well-being, housing and environment, educational well-being, health and safety, risk behaviours, and quality of school life (Table 3). Their main advantage is that they are standardised for many countries and regularly collected. However, they might be difficult to apply at parents/family and child-level, and also for some specialised or alternative care services, for example if data is collected through the household or school.

Source: P4EC CEE/CIS Consultancy Group, 2013.

---

12 PWC (2010) Valuing social impacts: Should government pay for results?
When outcomes cannot be measured, it is often appropriate to focus measurement on outputs. Outputs are intermediate steps along the transformation chain of resources, the results of social services provision that can be clearly stated or measured using agreed or available indicators. Examples of links between outputs and outcomes from Malta are shown in Table 4. It is not necessary for all identified services to specify outcomes/outputs at all four levels in the framework (Figure 2): society, community, parents/family, and individual child. Rather a hierarchy of outcomes/outputs should be clearly set out.

---

1. Policy relevance: High: governments can directly intervene with the family or individual through established policies, or through multiple secondary interventions. Medium: government relies on third-party intervention (professional or community [non-familial] actors). Low: no established routes for government intervention. In practice, no “low” policy relevant indicators were retained. An example of such an indicator might be, for example, peer relationships.

2. Belgian data is for 1997.


"x" refers to where selection criteria for the indicator or dimension are met; 
"*x" refers to where selection criteria for the indicator or dimension are not well met.

Source: OECD, 2009

---

Table 4: Malta pilot example of outputs and outcomes

<table>
<thead>
<tr>
<th>Social services or programmes to support children and families</th>
<th>Outputs</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal Unemployment assistance</td>
<td>Number of people received cash transfers</td>
<td>Maintain living standards Ensure income security</td>
</tr>
<tr>
<td>Targeted Children’s Allowance</td>
<td>Number of families receiving Children’s Allowance</td>
<td>Decrease level of vulnerability for poor families with children</td>
</tr>
<tr>
<td>Specialised Child Protection Services</td>
<td>Number of child protection cases</td>
<td>Strengthened family Preventing family separation</td>
</tr>
<tr>
<td>Highly specialised Ghabex shelter – mothers (victims of domestic violence) and children</td>
<td>Number of consultations/training provided to clients</td>
<td>Better safeguarding Lower level of injuries Lower level of significant harm</td>
</tr>
<tr>
<td>Alternative care High support service – ensuring children in residential care acquire independent living skills</td>
<td>Number of children in residential care</td>
<td>Improved skills and competencies for independent life</td>
</tr>
</tbody>
</table>


2.3 Combining costs and outcomes

At this point in the inquiry, it becomes necessary to explore the available methodological approaches that use information on costs and outcomes to inform a judgement on the value of a given service. A useful typology of evaluation approaches, originating from the health sector, considers two essential questions (Table 5):

- Does the methodology allow the **comparison** of two or more alternatives?
- Does the methodology use information on both **costs and outcomes**?

The methodologies in the bottom right corner of Table 5 (e.g. cost-effectiveness analysis, cost-benefit analysis), which compare two or more alternatives using both costs and outcomes, are considered **full evaluation methodologies**; the others are **partial evaluation methodologies**. The main features of the full evaluation methodologies are outlined below.

- **Cost-effectiveness analysis (CEA)** assesses the cost of the intervention against its ‘effectiveness’, defined by measures of outputs or outcomes. Outcomes are usually ‘natural units’ such as lives saved, number of family separations prevented, etc. In principle, CEA offers a means by which services can be compared to similar services that deliver the same outputs or outcomes.
- **Cost–benefit analysis (CBA)** compares the cost of the intervention and economic benefits expressed in monetary terms. Monetary returns can be estimated to the commissioner of the services (e.g. a public sector agency), participants, or society as a whole.
- **Cost–utility analysis** is similar to CEA, but the outcomes of the service are translated into a standard measure of economic ‘utility’ – the quality adjusted life year (QALY). Its most common application is in health economics, where the benefits of an intervention can be

---

16 Save the Children (2014) Child Protection Outcome Indicators.
expressed in terms of QALYs, allowing for more consistent comparisons across a wide range of interventions acting on different aspects of health and well-being.\(^{17}\)

- **CCA** is a form of CEA which presents the range of benefits identified alongside costs incurred without aggregating them into a single metric (e.g. a cost-effectiveness ratio), and instead leaves the decision makers to incorporate their own considerations when judging the relative merits of the intervention or programme.

### Table 5: Types of economic evaluation methodology

<table>
<thead>
<tr>
<th>Are two or more alternatives being compared?</th>
<th>Are costs and outcomes considered simultaneously?</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Only costs</td>
<td>Cost of service</td>
</tr>
<tr>
<td>Only outcomes</td>
<td>Outcome description</td>
</tr>
<tr>
<td>Yes</td>
<td>Cost analysis</td>
</tr>
<tr>
<td></td>
<td>Outcome (efficacy) analysis</td>
</tr>
<tr>
<td></td>
<td>Cost-consequence analysis</td>
</tr>
<tr>
<td></td>
<td>Cost-effectiveness analysis</td>
</tr>
<tr>
<td></td>
<td>Cost-utility analysis</td>
</tr>
<tr>
<td></td>
<td>Cost-benefit analysis</td>
</tr>
</tbody>
</table>

Source: Adapted from Drummond *et al.* (2006) Methods of economic evaluation in healthcare.

The scope of *Childonomics*\(^ {18}\) includes both reflecting on the ‘long-term social and economic return of investing in children and families’ as well as consideration of ‘different types of costs of different services and approaches to supporting children and families in vulnerable situations… and links them to the expected outcomes of using these services’. It is, as such, apparent that a **full evaluation methodology** is appropriate in this case given the necessity to enable comparisons across services and to incorporate both costs and outcomes.

It is essential, however, to consider further that the methodological approach to be developed as part of *Childonomics* is deliberately broad, aiming to be ‘as comprehensive as possible, addressing the interplay between social welfare, child protection, health and education services’.\(^ {19}\) As such, it is necessary to also consider methodologies that can capture some of the specificities of these sectors and permit comparisons of investments and outcomes related to different parts of the broader social services system.

**Return on Investment (ROI)** is a methodology originally developed to evaluate training and performance improvement programmes, but its application has been extended to the evaluation of any type of programme. It bears many similarities with CBA in key steps such as converting benefits to monetary value, isolating benefits that would have happened anyway, and comparing monetary benefits to programme costs.\(^ {20}\) The difference is that CBA usually returns results either as a **benefit–cost ratio** (e.g. a ratio of 2.5:1 suggests that for every $1 spent the programme generates $2.5 worth of value) or as a **total net benefit** (e.g. a $50,000 net programme benefit

---


\(^{19}\) Ibid.

may result from $200,000 worth of benefits minus $150,000 programme costs), while ROI returns a **percentage** (e.g. a 150% ROI suggests that for every $1 spent in the program, $1.5 worth of benefits are returned after costs). The ROI approach can be used to evaluate public sector services; however, its use in the social sector has been limited to date.

**SROI** is a framework for understanding, measuring, and managing the outcomes (and impact) of an organisation’s activities or services. It is designed to measure change in ways that are relevant to the people or organisations that experience or contribute to this change. It analyses how change is created by measuring social, environmental, and economic outcomes, and uses monetary values to represent them.

There are two types of SROI analysis:

- **Evaluative** SROIs are conducted retrospectively and are based on outcomes that have already taken place.
- **Forecast** SROIs predict how much social value will be created if the activities meet their intended outcomes. This can be particularly helpful at the planning stage of a project.

By placing a monetary value on outcomes, they can be added up and compared with the investment made. This results in a ratio of total benefits to total investment. While the ratio is important, the SROI process is designed to present a **story of change**, including both qualitative and quantitative findings, and provides information to help organisations maximise their outcomes. The SROI can also provide useful information for investors and commissioners. However, the ratio should be used with caution, as it is not intended to provide a comparison with other organisations.

Finally, a **Value for Money (VfM)** analysis is an approach for measuring the value of programmes which relies on the concept of a results chain (or ‘transformation chain’); this follows the transformation of monetary resources into programme outputs, which are converted through process in outputs and outcomes towards generating impact (Figure 6). While VfM concentrates on the relationship between inputs and outcomes-impact, VfM can and should be measured at all stages of the results chain. There are two essential considerations for conducting a credible VfM: having a solid, logical results chain at the foundation of the programme design and evaluation; and contextualising the result by being aware and explicit about the social, institutional, and economic realities (and other aspects) in which the evaluation is conducted. There are examples of the use of VfM to evaluate child programmes, such as children’s centre services in England.22

---

Figure 6: VfM 3E framework applied to social protection systems


Overall, the joint evaluation of costs and benefits for child programmes is a relatively new area and few studies are available.\(^{23}\) Most of them focus on programme costs and cost per beneficiary (a measure of cost-efficiency), and less on outcomes (cost-effectiveness). For example, a review of evaluations of interventions for children and young people with speech, language, and communication needs pointed out the paucity of available studies combining costs and outcomes.\(^{24}\) There are also ambiguous uses of the methodologies. The obvious candidate here is CEA, which is often used to denote cost studies with little explicit consideration for outcomes; at most, these are cost-efficiency studies if they calculate the unit cost per programme beneficiary.

However, the interest towards developing more sophisticated tools and approaches for assessing the value of social services appears to be increasing. One example of such a tool is the New Zealand’s Government Guide to producing a social investment evidence brief (2017), which combines ‘financial cost and benefits information with information on effectiveness’ to determine what works, for whom, and at what cost in terms of social interventions.\(^{25}\) The specific purpose of the social investment evidence brief is to inform decisions on which initiatives to be piloted, tried, and

---


\(^{24}\) Law J., Beecham J. and Lindsay G. (2010) Effectiveness, costing and cost-effectiveness of interventions for children and young people with speech, language and communication needs. Department for Education.

and evaluated in the New Zealand context based on previous experiences elsewhere. Essentially, this 10-step guide helps users to formulate the question, search for available evidence on costs and benefits, rate the evidence using the Maryland Scientific Methods Scale, and write up the results. The outputs of such an exercise, aimed at being produced in a short timeframe (four to six weeks) suitable to inform policy decisions, can be coupled with other tools. In this context, of utmost interest is the Social Investment Analytical Layer, which structures and harmonises information on a given individual’s contact with various government agencies (e.g. health facilities, education facilities, welfare institutions) across their lifetime. This allows government agencies to better understand where they can make investments in order to improve their service.

In the same vein, the approach developed by Wulczyn et al. (2009) at the University of Chicago is tailored to child services, specifically to understanding better the ROI for child welfare agencies. The authors proposed a framework that distinguishes between several types of outcome (mission-critical, safety, and permanency), and process and quality as key dimensions to service provision that can be evaluated and directly determine outcomes. While not quantitative in nature, this approach is predicated on making better use of routinely collected data in order to monitor the performance of services.

3 Proposed methodology

This section outlines the proposed evaluation methodology to assess the economic return of investment into child services and interventions. It is essential to establish at the outset what the methodology is and what is not—or at least what it aims and does not aim to be.

The role of the methodology is to guide researchers and practitioners towards gathering evidence iteratively in consultation with stakeholders in order to formulate a judgement on the comparative value of services, programmes, or policies based on information on their costs and outcomes. The methodology comprises principles, methods, and reporting standards the application of which guides its users towards producing relevant, robust, and comparable results. In this sense, the methodology is somewhat similar to the concept of a ‘reference case’, which has gained some traction in the global health arena. One such example is the Gates Reference Case for conducting economic evaluations in low- and middle-income countries (Box 1). While the scope of Childonomics is not to produce a reference case for child services—a chief reason being that it is not centred on the uniform application of a tried and tested evaluation methodology in the sector—it shares some of the aspirations of a reference case, e.g. optimising the value of economic information for decision making in the social sector, as well as improving transparency and clarity.

Box 1: The Gates Reference Case

The Gates Reference Case is a set of principles, methodological standards, and reporting standards to support health economic evaluations funded by the Bill and Melinda Gates Foundation. It is intended to support:

- the routine application of certain fundamental principles by researchers and decision makers in order to optimise the value of economic evaluation in informing good decisions in health;
- the use of methods that adhere to the same fundamental principles to facilitate comparison of the quality and relevance of economic evaluations when used to inform decisions in different contexts;
- a minimum standard of methodological quality to ensure that economic evaluations are fit for purpose; and
- minimum reporting standards to ensure clarity and transparency of economic evaluations, and to improve the comparability of both the content and results in different contexts.

The reference case should be considered an aid to thinking, by offering a systematic framework of assessment for making decisions. It need not be applied inflexibly; rather, it should be used to optimise the use of specific methods and existing evidence to produce useful and high-quality analysis.


The methodology does not offer a menu of the ‘best’ analytical choices— for example, which costs to incorporate or which outcome indicators are ideal to measure societal benefit. These would be determined on a case-by-case basis, depending on the nature of the question, the setting, and the final beneficiaries of the methodology. The methodology offers, however, some key principles and a menu of considerations to account for.

The methodology’s focus on establishing the ‘relative value of services, programmes and policies’ is not to be interpreted as meaning that it can be used as a tool for conducting an audit of available services. It cannot, and is not meant to, produce judgements on whether existing services are functioning appropriately, e.g. whether resources are being utilised efficiently and effectively, in that they reach the appropriate beneficiaries, to a certain standard etc. Neither does
it aim to suggest explicit reconfigurations in service financing, organisation, and delivery. As such, it is not a normative analysis of ‘what should be in place’, as compared to ‘what is currently in place’. Instead, it focuses on the (expected) return of ‘what is (or is likely to be) in place’.

However, as data on standards, resources, service users, and activities are collected and analysed throughout the process, the users of the methodology may inevitably develop views on some of the aspects outlined above. It must be borne in mind that other than the rights-based approach underpinning the conceptual framework (which guides the types of outcomes and classification of child services), the methodology does not impose or advocate any particular service standards; neither does it invite its users to populate it with their own interpretations of the context. The series of matrices accompanying the methodology as tools to support its application are to be populated with factual data.

Similarly, the methodology is not a field manual; given the range of topics it includes (e.g. cost estimation, stakeholder consultation, mathematical modelling, evidence synthesis, etc.), it is beyond the scope of this paper to be or to include a primer for applying such methods. We refer readers and potential users throughout to resources that offer more in-depth guidance on selected aspects.

### 3.1 The Childonomics approach

The methodology borrows elements from SROI and other approaches (see previous section page 12-13) in terms of estimating costs and outcomes; however, it is neither of these in a pure sense because of the particularities of Childonomics:

- **It is not meant to be purely SROI** because SROI is designed ‘as a practical management tool that can be used by both small and large organisations, rather than from a macro perspective’.\(^\text{27}\) **Childonomics** focuses on services and policies, not on organisations. It inherently takes a broad evaluative perspective, especially about the outcomes it considers, and aims to allow comparisons between services and policies. Hence, it arguably has a broader scope than SROI can provide.

- **It is not meant to be purely CEA or CBA** because of anticipated difficulties with aggregating into a single measure all relevant outcomes (economic, social, health, education, psychological) resulting from a service, be it a natural unit such as ‘life saved’ (CEA), a monetary unit such as ‘net monetary gain’ or ‘ROI ratio’ (CBA). One of the key aims of **Childonomics** is to communicate benefits and value beyond a single number, and to reflect a broad range of outcomes and their interplay.

The **full evaluation methodology** that appears to be closest to the aims of **Childonomics** is **CCA**, which presents costs and outcomes side by side in a disaggregated manner. CCA is a form of CEA which presents the range of benefits identified alongside costs incurred without aggregating them in a single metric (e.g. a cost-effectiveness ratio), instead leaving the decision makers (and users of the methodology, in a broader sense) to incorporate their own considerations when judging the relative merits of the intervention or programme. CCA is particularly valuable when different outcomes cannot be incorporated into an index measure, as is the case here.

The methodology allows the comparison between a comparator, usually a policy or a programme that is already being implemented, with one or more alternative(s) which are either not being

---

delivered at the time of the evaluation, but could be introduced, or are delivered but require scaling-up or some adjustments. In principle, there is no limit to the number of alternatives included in the analysis; however, it is important to keep the comparison manageable. With one comparator and more than three alternatives, the amount of information can become difficult to synthesise. Therefore, it is recommended to focus on one or two well-chosen alternatives.

While the methodology aims to identify and make use of the available evidence concerning the interventions being analysed, it is essential to note that their expected costs and outcomes are likely to be influenced by the political dimension of the environment. Irrespective of how strongly the available evidence may favour one course of action or the other, it may be the case that the policy discourse or priorities of the moment do not favour its uptake or successful implementation. No amount of analysis can compensate for such influences. As such, the methodology assumes that its users consider the broader context at the outset and anchor the evaluation in such realities (Step 0).

The logic of the methodology is as follows, as set out in Figure 7:

- Firstly, an assessment of the policy context informs the delineation of the analytical scope and the purpose of the CCA: this includes a description of the comparator and its alternative(s), and a specification of the time horizon and of the economic perspective for measuring costs.
- Outcomes are then selected and specified using an outcome matrix.
- Costs and outcomes are then measured for both the comparator and the alternative(s).
- Finally, incremental quantities (net impacts) are calculated as differences between comparator and alternative(s).
- Results are presented in a results matrix together with an indication of the strength of evidence supporting them.

Given the intersectoral nature of the methodology, its application would ideally rely on a multi-disciplinary team whose members’ skills reflect its nature. At a minimum, a child services/social work specialist and an economist would form the core of the team. Depending on the scope of the analysis, the specific approaches taken, available data, and resources available for the analysis, they may be supported by specialists in other fields (e.g. statistics, demography, public financial management, public sector governance, child participation, human rights etc.).

The layout of the methodology guides how its application is to be reported, with the outcomes matrix, results matrix, and summative narrative forming the eventual final deliverables synthesising the analysis for consideration by the target audience.
### Figure 7: Methodology outline

<table>
<thead>
<tr>
<th>Steps</th>
<th>Key activities</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conduct analysis</td>
<td>Stakeholder consultation</td>
<td>Outcome matrix</td>
</tr>
<tr>
<td>Assess context</td>
<td>Policy analysis</td>
<td></td>
</tr>
<tr>
<td>Establish scope</td>
<td>Literature review</td>
<td></td>
</tr>
<tr>
<td>Specify and measure outcomes</td>
<td>Stakeholder consultation</td>
<td>Results matrix</td>
</tr>
<tr>
<td>Specify and measure costs</td>
<td>Literature review(s)</td>
<td></td>
</tr>
<tr>
<td>Estimate net costs and outcomes</td>
<td>Primary data collection Economic modelling Other modelling forms</td>
<td>Summative narrative</td>
</tr>
<tr>
<td>Assess strength of evidence</td>
<td>Validation workshop</td>
<td></td>
</tr>
<tr>
<td>Construct narrative</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Authors
The steps are outlined in more detail below.

### 3.2 Analysis

#### 3.2.1 Step 0: Assess context

An analysis of the value of child services in any setting requires situating it in its respective policy context in order to ensure that the main question is relevant from the outset and that the findings of the analysis can inform action. For the users of the methodology, this entails formalising their initial understanding of the existing child policy objectives of the menu of available programmes or interventions, and of the political climate towards the social sector in general and child services in particular. The findings of the context assessment will guide further choices relating to the purpose and scope of applying the methodology (Step 1 below).

The following questions can guide the users of the methodology in this initial assessment:

1. What outcomes for children are important for the key stakeholders – Government, NGO, academic community?
2. What programmes and services have been/are successful in achieving these outcomes?
3. Which services are considered as the priorities for development and funding and why? Supplemental areas may also be explored, such as: what factors underpin the process of introducing or withdrawing programmes and services? How do intersectoral or inter-ministerial cooperation affect decisions on money allocation or reallocation according to needs? What are the levels of governance involved and the competencies, responsibilities, and accountability of decision makers?
4. What are the core inputs and activities of priority programmes and services? For example, what are the costs of social workers visiting families, of counselling, psychological accompaniment, training/education, income maximisation or of fulfilling social administration tasks. Are there critical aspects to how services are delivered that impact on their effectiveness, e.g. working culture and core values, application of basic standards?
5. What are the available data sources to inform answers to questions 2–4?

The questions above can be answered by applying a mix of approaches, which include:

- Conducting stakeholder consultations with key policymakers at national and local level, and with other national and local authority stakeholders;
- Reviewing policy documents, e.g. legislation, position papers, policies, strategies; and
- Reviewing published literature, e.g. academic papers, NGO reports.

#### 3.2.2 Step 1: Establish scope

This initial step entails clarifying the boundaries of the analysis, namely making decisions and being able to answer the following questions clearly:

- **What is the main purpose of the study and who is the main audience?** The analysis could be used to inform decisions on policy or resource (re)allocation, to support advocacy or to contribute to the research body in a specific area. As such, the audience may include the organisation(s) delivering the services being compared, the state- or national-level policy makers, donors, NGOs, researchers or others. Different audiences are interested in different types of question, work with different types of costs, outcomes, and time horizons; therefore,
clarifying this aspect at the outset can ensure the scope is clearly defined and facilitate the process of validating findings and constructing the summative narrative.

- **Which services or types of service are being compared?** This entails describing the comparator and the alternative intervention(s) in terms of the activities they comprise (e.g. counselling and cash transfers for families), the characteristics of the target population (e.g. children aged five to nine years), and the setting (e.g. local, regional or national). The comparator and alternative interventions should be contributing towards the same or similar outcomes for children, families, communities, and society. If common outcomes cannot be defined, then they cannot be compared. In this regard the process for identifying the services being compared and specifying outcomes in Step 2 below may be iterative: initial assumptions about the comparability of services are tested in Step 2; then Step 1 may need to be re-applied to adjust the comparators if common outcomes cannot be specified.

- **Is the analysis prospective or retrospective?** This influences the data and methods that can be applied. A retrospective analysis (i.e. of services/policies that are already in place) will rely heavily on secondary data (e.g. existing research and routine data monitoring systems) to inform the realisation of the outcomes. A prospective analysis (i.e. of services/policies that have not yet been introduced in the given setting, or a forecast) will often involve data available from other settings, and may require de novo data collection to be conducted.

- **What is the economic perspective of the analysis?** This entails being explicit about who incurs the costs, and what types of cost are included in and excluded from the analysis. The analysis can be conducted from several perspectives and results presented jointly – the only pre-condition is being clear about what the perspectives are. For example, the analysis can take a financial perspective where only costs incurred as monetary transfers are considered, e.g. costs incurred by institutional service providers such as district child protection authorities. As an alternative, or as a complement to the financial perspective, a social/economic perspective can be taken to consider broader costs to society which are not necessarily incurred as financial transfers, for example, accounting for the value of lost labour productivity due to sickness, caring for a family member, or adverse events associated with an intervention.

- **What is the time horizon for considering costs and effects?** From a comparative perspective, it makes sense to have a time horizon that accounts for all expected differences in costs and outcomes between the comparator and the alternative(s). As such, the choice of analytical perspective and outcomes (Steps 2 and 3 below) are key determinants of the choice of time horizon. Ideally, this decision should be informed by evidence on how the outcomes of interest evolve over time. It is not good practice to shorten the time horizon purely for analytical convenience – this introduces the risk of mis-estimating results. At the same time, it is not analytically efficient to take a time horizon longer than expected differences between alternatives – this may introduce noise into the estimation.

### 3.2.3 Step 2: Specify outcomes at each level of the Childonomics framework

In the Childonomics framework, outcomes are specified at four levels: individual, parents/family, community, and society (Figure 2). The levels are not, however, prescriptive in determining which outcomes and indicators are ‘appropriate’ or ‘best’. The choice will depend on the specific question that the analysis is attempting to answer. In order to identify the outcomes and their associated indicators, there are two perspectives to consider:

- **A data-informed perspective**, which involves using available datasets, consulting official and (routinely collected) statistics on service use, conducting literature reviews, etc.
- **A beneficiary-informed perspective**, which entails engaging service beneficiaries (e.g. children and their families) and other stakeholders in structured dialogue, revealing what matters to them. In a broader sense, it is essential to engage all relevant stakeholders in this process – an element that the SROI methodology also features strongly (Step 1.2: Identifying stakeholders, and Step 1.3: Deciding how to involve stakeholders).

The child rights-informed approach of *Childonomics* set out in the conceptual narrative (Bilson *et al.*, 2017) similarly does not predetermine the outcomes of value at each level of the framework. International child rights frameworks and guidance offers a lens for approaching the selection of outcomes and indicators, rather than for selection itself. The methodology rather emphasises the agency of stakeholders in assigning value to outcomes and identifying meaningful indicators. In particular, it gives a central place to the right of the child to be heard in decisions affecting them (UNCRC, Article 12).

Powell and Smith (2009, in Douglas, 2011)28 recognise that children’s right to participate may be compromised if we project ‘vulnerability’ onto children and limit their participation. Douglas (2011: 101)29 goes on to say that if this happens, it is in direct violation of their rights to have an opinion and to have those opinions taken seriously. It is recognised that there are difficulties and nuances in consulting service users, especially children, children in care and care-leavers, but the investment it would take in order to consult with children ensures the added value of their participation.

It is important to include the various groups of service users: to exclude them is to exclude a key group of stakeholders and a key source of knowledge and information in the community. The research and the research outcomes will be better informed if service users are facilitated as equal and active participants. Their insights represent an embodied experience of system design and their participation can actively speak to the lived reality of how the system affects them, as well as how and whether the services are meeting identified needs. Their contributions can also speak to the direct or indirect costs that are incurred both within and outside the services, to compensate for a lack of services or in order to ensure service access.

There is no hard and fast process for selecting the outcomes and there is no ‘core’ set of outcomes. The usual process would involve reviewing available theory, data, and research outputs, as well as consulting service beneficiaries and other stakeholders, in what is likely to be a deliberative, iterative process guided by the scope of the analysis (Step 1). This will result in completing an **outcome matrix**, of which an illustrative example is presented in Table 6. The outcome matrix summarises the outcome and output domains, specific indicators, and potential data sources for their measurement.

---


29 Ibid.
### Table 6: Illustrative outcome matrix from Romania pilot (day centres for children with disabilities)

<table>
<thead>
<tr>
<th>Domain</th>
<th>Indicator</th>
<th>Data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outputs for service users</td>
<td>Child development</td>
<td>Age and ability-appropriate developmental milestones (including ECD) met, progress in communication, social skills and behaviour</td>
</tr>
<tr>
<td></td>
<td>Increased capacity of parents to provide care</td>
<td>Parent knowledge and competency</td>
</tr>
<tr>
<td>Child</td>
<td>Children prepared for school/education</td>
<td>Literacy, progress in education during and after service use</td>
</tr>
<tr>
<td>Family</td>
<td>Parents employed and relationships stable</td>
<td>Household income and structure</td>
</tr>
<tr>
<td>Community</td>
<td>Reduced separation of children from parents</td>
<td>Number of children with disabilities from communities targeted by day centres entering long-term alternative care</td>
</tr>
<tr>
<td>Society</td>
<td>Increased education and inclusion of children with disabilities</td>
<td>AROPE disaggregated for children with disabilities from different services at different stages of adulthood</td>
</tr>
</tbody>
</table>

Source: Authors, based on document review and interviews with staff and service users at two day centres for children with disabilities in Romania, 2017.30

### Ideally, outcome domains and indicators would be identified for all four outcome levels, from individual to societal. This may not always be possible with the available data. To preserve consistency across analyses using the Childonomics framework, empty cells in the outcomes matrix will need be justified in terms of why indicators at a certain level or suitable indicators for them could not be identified or if they are not relevant for the intervention(s) or research question at hand.

#### 3.2.4 Step 3: Measure outcomes

Once the outcomes are specified, values for each outcome indicator and for each intervention (the comparator and the alternatives) will be informed by various types of data. Some form of modelling may be necessary, where modelling is understood here in a broad sense as a process comprising several generic steps:

---

30 AROPE is the ‘at risk of poverty and exclusion rate’, an indicator used by the European Union in monitoring the implementation of 2020 strategies in EU member states.
• **Identifying the relevant data**, understood as obtaining information on the target outcome indicators. Depending on the indicator, data can be either quantitative (e.g. poverty rates) or qualitative (e.g. perceived experience of using the service). There may be cases where data are already available, for example in databases, published studies, or reviews of the literature. In other cases, data may have to be collected as part of the analysis.

• **Assessing the extent to which data constitutes evidence for the said outcomes.** This involves making a judgement on **validity and generalisability**. What these concepts mean depends on the nature of the relevant data.

For quantitative research, internal validity essentially refers to the certainty with and extent to which observed effects are due to having been exposed to the intervention. Is there a convincing causal relationship between the intervention and the effect? Are other factors at play that may determine outcomes? For example, how strongly do the data support the hypothesis that institutional placement between ages five and nine years increases the risk of poverty in adult life? Additionally, there is the question of incorporating appropriately the size of this effect. The SROI methodology uses several concepts that are also relevant here (Box 2).

For qualitative research, validity refers to the appropriateness of the tools, processes, and data to answer the question at hand. For example, was the chosen methodology appropriate? Was the study design appropriate for the methodology? Was the sampling strategy and analysis method appropriate for the study design? And are the results and conclusions valid for the sample and context?

**Box 2: Incorporating effect size**

<table>
<thead>
<tr>
<th>Attribution</th>
<th>An assessment of how much of the observed outcome was caused by the contribution of other services, programmes, policies, or organisations.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deadweight</td>
<td>A measure of the amount of outcome that would have happened even if the activity had not taken place.</td>
</tr>
<tr>
<td>Displacement</td>
<td>An assessment of how much of the outcome has displaced other outcomes.</td>
</tr>
<tr>
<td>Drop-off</td>
<td>The deterioration of an outcome over time.</td>
</tr>
<tr>
<td>Impact</td>
<td>The difference between the outcome for participants, taking into account what would have happened anyway, the contribution of others, and the length of time the outcomes last.</td>
</tr>
</tbody>
</table>


For **generalisability (external validity)**, the focus is on determining the alignment between how and where data were obtained and the scope of the present analysis. An example of this may be the extent to which findings from another country or continent are likely to hold in the respective setting. For example, are effects observed in South Asia likely to hold in Eastern Europe? Or are effects observed in marginalised populations likely to hold in the general population?

We return to these issues in the sub-section on ‘Strength of evidence’ below. What is essential here is for users of the methodology to reflect on these issues in relation to the data available to them. This will determine at an early stage which data are to be retained for further analysis and which are to be discarded. This is particularly important when a decision is made that a distinct data collection exercise needs to be conducted as part of the methodology, e.g. conducting a small survey or focus group discussions. Such exercises may be time-consuming and costly, therefore anticipating the extent to which data that are to be collected can contribute to the evidence case is strongly encouraged.
• Potentially processing the data, e.g. by combining data from different sources or of different types to arrive at new insights. For quantitative data, this is often done by specifying mathematical relationships. One example is conducting a meta-analysis, i.e. aggregating intervention effects from several comparable studies to obtain an average intervention effect. Another example is the possibility of combining a demographic model which projects future numbers of the target population, e.g. young people aged 15–18, with socio-economic status to predict absolute numbers of people at risk of poverty. For qualitative data, triangulation is an example of a well-used social science approach that brings together several types of data sources (or collection methods, or analysts, or theories) to create a richer, more rounded understanding of the issue at hand.

It is beyond the scope of this paper to specify which methodologies and methods to use in which situations or for which types of questions. We summarise in Box 3 the main approaches and data sources we used during the pilot work in Malta and Romania. This list is by no means exhaustive – the pilot report contains more detail and an annex of instruments and methods used.

**Box 3: Summary of approaches and data sources for outcome specification and estimation used in Malta and Romania**

**Primary data sources**

**Key informant interviews:** National government specialists and decision makers; service managers; social workers; social pedagogues; case managers.

**Case studies:** Family support case reviews; interviews with care leavers and young people in care.

**Service user consultations:** Children in residential care; children in foster care; foster carers; parents.

**Secondary data sources**

- service statistics;
- review of child protection and service delivery database parameters;
- service manuals/procedures/standards; and
- policy analysis and service provision mapping.

3.2.5 **Step 4: Calculate costs**

Ideally, three types of costs are calculated:

- **total costs** of delivering the service/implementing the policy scenario;
- **total average costs per beneficiary** (child, family etc.); and
- **marginal cost per beneficiary**, i.e. cost of delivering the service to one additional beneficiary (child, family etc.) once the overhead and set-up costs have been accounted for.

These three types of cost indicators have been proposed because of their relevance for the scope of *Childonomics* and applicability to most types of intervention. They constitute an indicative core list of indicators. Specific applications can also consider other types of cost, as appropriate. For example, a recurrent, high-intensity service can examine the total cost per beneficiary per month or per intervention. For some services, a small proportion of users are highly resource-intensive compared to the rest and these incur a large proportion of the total costs, in which case the average cost per beneficiary may be misleading. The median cost or the range of costs can also be presented in such cases, bearing in mind, however, that they have a more limited application for planning purposes compared to the arithmetic mean.
How to go about costing the interventions depends on the setting and the economic perspective of the study (Step 1 ‘Establish scope’). For the **direct costs of service provision**, several approaches can be taken. The comparator is usually an intervention already implemented, therefore a top-down costing approach informed by expenditure data, e.g. from child protection authorities, can be used to inform the estimates. For alternative interventions that have not yet been implemented widely, a bottom-up costing exercise may be required; this entails identifying and measuring the types of resources (staff, equipment, consumables, buildings etc.) and assigning them monetary values. Accounting for set-up costs is equally important: the building, equipment, recruitment of trained staff, initial capacity building, etc.

The valuation of **direct non-service related costs** (e.g. service access fees) and of **indirect costs** (e.g. lost productivity of beneficiaries, psychological counselling) usually involves a bottom-up calculation informed by the number of beneficiaries. The results of household surveys are usually used to inform these calculations, e.g. how many hours/days an individual spent off work. It is also important to be clear and consistent about the base year of cost estimation and apply appropriate treatment to costs, e.g. discounting to estimate the present value of future costs.

The interested reader is encouraged to consult specialised literature on conducting cost analyses in the social sector. The Guide to SROI\(^{31}\) and the PSSRU guide to cost estimation in social services\(^ {32}\) are two excellent examples.

### 3.2.6 Step 5: Calculate net costs and outcomes

In the final step, net costs and outcomes are determined as **differences for each indicator between alternative(s) and the comparator**. This operation may be more straightforward for quantitative indicators, such as costs: it involves subtracting respective quantities for the comparator from the quantities for the alternative interventions, e.g. the average incremental total programme cost is the difference between the total cost for the alternative programme and the total cost for the comparator programme.

For some outcomes, only a qualitative judgement can be made due to lack of appropriate measurements, e.g. the impact on future psychological well-being or self-esteem. In cases where the net impact cannot be quantified, a judgement can be made based on the available evidence, e.g. a given intervention has a beneficial effect on the sense of independence and self-confidence of its beneficiaries relative to the comparator.

### 3.3 Strength of evidence

**Strength of evidence** is a distinct consideration accompanying the net impacts as estimated at the end of Step 5 (above). In a field as complex as child services, determinations of costs and outcomes are not equally credible across evaluations. For some estimations, strong modelling assumptions are required (e.g. choosing distant proxy indicators for outcome measurement). For others, the current level of knowledge may not go beyond demonstrated association, which makes it difficult to say with certainty whether something ‘really works’ or not. In other cases, cultural generalisability can play a substantial role (e.g. certain models of care are more

---


© Oxford Policy Management
acceptable to beneficiaries in a given country than other models of care or than the same category of beneficiaries in another country).

A number of taxonomies allow grading the levels of evidence, of which we present several examples in Annex A. They focus on determining whether the intervention or programme works and their underlying principle is the same: the way in which the evidence is produced gives an indication of how strong the evidence is. For example, anecdotal evidence (e.g. testimonies of social workers indicate improvements in children’s emotional well-being) represents a much weaker indication of effectiveness compared to an experimental study (e.g. a randomised controlled trial with a positive, statistically significant improvement in self-confidence measured using a standardised instrument).

Strength of evidence is not to be confused with effect size. It may be the case that a relatively small effect size is supported by robust evidence, e.g. a systematic review of randomised controlled trials indicates an average, statistically significant 2% reduction of school drop-out in children from families benefitting from counselling. In this hypothetical case, we can be fairly certain (subject to the Randomised Controlled Trials being of good quality) that counselling leads to a reduction of school drop-out rates, albeit of rather small magnitude. Conversely, an apparently significant effect size may be supported by weak evidence, e.g. anecdotal evidence from case workers points to a 50% reduction in family separation following the introduction of a high-intensity home visit service. It is very difficult to make causal claims on these grounds alone however because other factors may be at play.

An assessment of the strength of evidence depends on a number of factors, including what the evidence is being used for and who is interpreting it. Policymakers may have different interpretations and requirements from researchers for what constitutes good evidence. This aspect reiterates the importance of knowing the main audience for the findings of the analysis (step 0).

A slightly more nuanced concept is the evidence journey, which refers to the fact that evidence gathering is an iterative, continuous process rather than a straight-line race for a definitive answer. Alongside this continuum, an intervention, a programme or a policy may appear beneficial when it is not yet fully evidence-based, or (seemingly) contradictory pieces of evidence can arise. In this case, it is essential to identify where on this evidence journey a given intervention or programme can be situated. A useful framework for the placing of available information on an intervention/programme on the evidence journey is represented by the standards of evidence used for assessing intervention effectiveness developed by the Social Research Unit at Dartington, which considers four dimensions: evaluation quality; intervention impact; intervention specificity; and system readiness (Annex A.4). For each dimension, criteria differentiate between ‘good enough’ evidence and ‘best’ evidence.

To our knowledge, there is no ready-made taxonomy for strength of evidence that can be applied for this particular methodology. A complicating factor is that not all data are the results of a research design. Indeed, the values for some outcome indicators can result from a combination of methods for which no taxonomy is available. For example, results from a statistical analysis of hospitalisation data can be coupled with a demographic model to estimate the future number of children requiring hospitalisation for drug abuse.

We propose here a simple taxonomy for each net impact (cost or outcome) based on the two considerations referred to earlier in Step 3, namely **validity** and **generalisability**. Each impact estimate is categorised as **High**, **Moderate** or **Low** for both validity and generalisability. The overall **strength of evidence** combines assessments of the two (Table 7). The definitions of the assessment categories High/Moderate/Low are outlined in Table 8, based on the Petticrew and Roberts (2003) framework (Annex A.3). We propose this framework because of its comprehensiveness in terms of mapping how appropriate various research designs, quantitative and qualitative, are for various types of questions and outcomes.

**Table 7: Summary of strength of evidence categories**

<table>
<thead>
<tr>
<th>Validity</th>
<th>Generalisability</th>
<th>Strength of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>High</td>
<td>Strong</td>
</tr>
<tr>
<td>High</td>
<td>Moderate</td>
<td>Strong</td>
</tr>
<tr>
<td>High</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Moderate</td>
<td>High</td>
<td>Strong</td>
</tr>
<tr>
<td>Moderate</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Moderate</td>
<td>Low</td>
<td>Indicative</td>
</tr>
<tr>
<td>Low</td>
<td>High</td>
<td>Moderate</td>
</tr>
<tr>
<td>Low</td>
<td>Moderate</td>
<td>Indicative</td>
</tr>
<tr>
<td>Low</td>
<td>Low</td>
<td>Indicative</td>
</tr>
</tbody>
</table>
Table 8: Definitions of strength of evidence categories

<table>
<thead>
<tr>
<th>Validity</th>
<th>Generalisability</th>
</tr>
</thead>
<tbody>
<tr>
<td>High – evidence is generated as a result of a research design that can be mapped to grade +++ in the Petticrew and Roberts framework.</td>
<td>High – net impacts are informed by context-specific data (e.g. the same country or region) and there are no reasons to suspect meaningful variations.</td>
</tr>
<tr>
<td>Moderate – evidence is generated as a result of a research design that can be mapped to grade ++ in the Petticrew and Roberts framework; or evidence is generated by combining results from several research products that can be mapped to grade ++ in the Petticrew and Roberts framework (e.g. decision modelling).</td>
<td>Moderate – net impacts are informed by data from a different context and there may be variations, but their magnitude is unlikely to influence the net impact (e.g. service cost data from a neighbouring county or country).</td>
</tr>
<tr>
<td>Low – evidence is generated as a result of a research design that can be mapped to grade + in the Petticrew and Roberts framework; or evidence is generated by combining results from several research products that can be mapped to grade + in the Petticrew and Roberts framework.</td>
<td>Low – net impacts are informed by evidence generated in a different context, variations likely exist and their impact on the net impact may be substantial.</td>
</tr>
</tbody>
</table>

3.4 The summative narrative: what do the findings mean and to whom?

NB: As set out in the rationale and outline of the CCA methodology in Section 3 of this paper, the following points should be re-emphasised at this stage when implementing the methodology:

The methodology’s focus is on establishing the relative value of services, programmes and policies. It is not a tool for conducting an audit of available services or producing judgements on whether existing services are functioning appropriately, e.g. whether resources are being utilised efficiently and effectively, in that they reach the appropriate beneficiaries promptly or to a given quality or standard. Neither does it aim to suggest explicit reconfigurations in service financing, organisation and delivery. The tool does not produce a normative analysis of what should be in place compared to what is currently in place. Instead, it focuses on the (expected) return of what is (likely to be) in place.

However, as data on standards, resources, service users, and activities are collected and analysed throughout the process, the users of the methodology may inevitably develop views on some of these aspects. It must be borne in mind that other than the rights-based approach underpinning the conceptual framework (which guides the types of outcomes and classification of child services), the methodology does not impose or advocate any particular service standards; neither does it invite its users to populate it with their own interpretations of the context. The series of matrices accompanying the methodology as tools to support its application are to be populated with factual data.

The results matrix presents (Table 9) jointly the net costs and outcomes, together with information on the strength of evidence. The results matrix offers both analysts and stakeholders a comprehensive view of all the findings. But how to interpret them? What do they actually say and how can this information be used?
### Table 9: Illustrative results matrix

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Domain</th>
<th>Indicator (examples)</th>
<th>Comparator – Family support</th>
<th>Alternative 1 – Residential care</th>
<th>Alternative 2 – Foster care</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Impact</td>
<td>Impact</td>
<td>Impact</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difference from comparator</td>
<td>Strength of evidence</td>
<td>Difference from comparator</td>
</tr>
<tr>
<td>Outputs</td>
<td>Outputs</td>
<td>Contact time per beneficiary</td>
<td></td>
<td>Strong</td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td>Child-level</td>
<td>Psychological well-being</td>
<td></td>
<td>Indicative</td>
<td>Indicative</td>
</tr>
<tr>
<td>Community</td>
<td>Number of prevented separations</td>
<td></td>
<td>Moderate</td>
<td></td>
<td>Moderate</td>
</tr>
<tr>
<td>Community</td>
<td>% variation in homelessness</td>
<td></td>
<td>Indicative</td>
<td></td>
<td>Indicative</td>
</tr>
<tr>
<td>Society</td>
<td>Productivity in adulthood Risk of poverty</td>
<td></td>
<td>Moderate Strong</td>
<td></td>
<td>Moderate Strong</td>
</tr>
<tr>
<td>Costs</td>
<td>Total cost</td>
<td>Total cost of programme per year</td>
<td></td>
<td>Strong</td>
<td></td>
</tr>
<tr>
<td>Costs</td>
<td>Total cost per beneficiary</td>
<td>Total cost per child Total cost per family</td>
<td></td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Costs</td>
<td>Marginal cost per beneficiary</td>
<td>Marginal cost per child Marginal cost per family</td>
<td></td>
<td>Moderate</td>
<td></td>
</tr>
</tbody>
</table>
The strength of evidence component (Section 3.3 above) attempted to support formulating an indication as to the credibility of findings; however, important unanswered questions remain:

- Are all outcomes equally important?
- How can the information in the results matrix be summarised, particularly on outcomes? Are there context-related considerations that may be affecting the information in the results matrix?
- What does the information in the results matrix say about the relative value of interventions/programmes/policies?
- What is worth doing or scaling-up and what is worth closer consideration?

Evaluation methodologies such as CBA, CEA, and SROI return a single metric (cost-benefit ratio, cost-effectiveness ratio, and ROI, respectively) that summarises information on the relative or intrinsic value of any intervention. A set of rules dictates how this figure can be interpreted. The cost-benefit ratio and ROI are sufficient, in and of themselves, to formulate a judgement given that 1 is the natural cut-off point: for values higher than 1 (e.g. a $2 return for $1 investment) the intervention is considered worth pursuing in and of itself, and conversely for values lower than 1. For cost-effectiveness ratios, comparing the specific cost-effectiveness ratio with a cost-effectiveness threshold (or range of thresholds) can help ascertain whether the intervention is cost-effective or not relative to its comparator. This is the approach taken in the UK by the National Institute for Health and Care Excellence, which has considered a health intervention unlikely to be cost-effective when it produces an additional QALY at an average additional cost (relative to the comparator) higher than £20,000–£30,000. Other health decision making bodies in Europe and across the world adopt a similar logic, of course with different threshold values.34

This is not to say that thresholds are the only way to incorporate cost-effectiveness information in decision making. Far from it – in healthcare, deliberative processes with varying degrees of transparency are used to combine considerations on effectiveness, safety, cost-effectiveness, budget impact, and others. Irrespective of the mechanism, the challenge (equally relevant for the social services sector) is how to merge scientific considerations relating to the evidence with social value considerations relating to ethics, preferences, and culture.35

The methodology proposed in this document is not amenable to such ‘hard’ judgement aids. Given the difficulty of reducing outcome valuation to a single metric and the built-in flexibility towards including a wide range of outcomes informed by various data sources, including beneficiaries’ views, no single ratio or threshold could be produced. This may appear as a limitation or an insurmountable challenge, and to an extent it is. However, it also provides an opportunity for stakeholders to remain meaningfully engaged throughout all the steps of the methodology, from defining the scope through to interpreting the results and deciding jointly what the results mean in their context.

We propose that the final step of the methodology is a validation workshop where the stakeholders, having been involved in the process from its early stages, are able to interact with the results. For some of them, this will be the first opportunity to engage with results concerning outcomes that usually lie outside their interests. As such, the workshop aims to construct a shared understanding of the results among those involved and use it as a foundation for result interpretation.

---

As such, the results matrix is not meant to be a static tool that returns a blunt answer. We rather envisage it as a tool that stakeholders engage with during the workshop by following a series of steps:

1. **Reviewing, understanding and validating the findings**: This would be conducted by the team responsible for conducting the analysis using Steps 1–5 outlined above, who present the purpose of the analysis, the analytical process, and the results matrix.

2. **Prioritising and validating the impacts**: This takes the form a guided discussion among the workshop participants on the priorities and strength of evidence, and reflection on the analysis of differences between the comparators.

3. **Reflecting on the results matrix and creation of a summative narrative**: This is a final group discussion, arriving at a joint conclusion on the relative value of the examined interventions and factors impacting their relative value based on the available evidence, stakeholder prioritisation, and interpretation of contextual issues identified by stakeholders. Reflection should consider not only the core research question and purpose of the analysis but also any unexpected results in terms of policy and decision making and further research that may be required.

The final output (summarised in Box 4) is the results matrix with a **summative narrative** that presents in a distilled form the findings of the analysis, the discussion from the validation workshop, and the final joint stakeholder and research team conclusion on the relative value of the examined interventions. The summative narrative provides an explanation of how the net outcomes for each alternative are valued by stakeholders and summarises the contextual issues, nuances of interpretation and feedback from stakeholders into a final (one-page) analysis that should always accompany the results matrix as the final output from the application of the Childonomics methodology.

The summative narrative should include key information from each step of the analysis, from establishing context and scope through to presenting and validating results, and taking into account a range of stakeholder perspectives.

**Box 4: Final output from the Childonomics CCA**

<table>
<thead>
<tr>
<th>Stakeholder-informed results matrix presenting comparative:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• measures of outputs and outcomes for child, family, community, and society that are valued by stakeholders; and</td>
</tr>
<tr>
<td>• costs – overall, per user, and marginal.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Summative narrative (one page) presenting considerations of:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• context;</td>
</tr>
<tr>
<td>• factors affecting, e.g. access, availability, quality and effectiveness; and</td>
</tr>
<tr>
<td>• stakeholder interpretations of the results and considerations of implications for decision making or next steps.</td>
</tr>
</tbody>
</table>
4 Final considerations

*Childonomics* is trying to do something very challenging and for which there are no existing standard models. Without reinventing the wheel, the *Childonomics* methodology draws on various established methods for appraising ROI and creates something slightly different that can help to put different types of investment and outcomes side by side for comparison.

The main challenges in implementing the methodology are primarily in defining outcomes which can be monitored while balancing complex value-laden questions of the relative ‘worth’ of particular outcomes. At the same time the methodology is dependent on the quality and validity of the data that is available and there will always be a challenge in ensuring there is strong evidence to be able to meaningfully attribute these outcomes to specific investments. The credibility of the findings is directly linked to the strength of evidence, which the methodology makes explicit specifically to allow a transparent examination.

The application of the methodology relies on rigorous evaluation and systematic collection of data from services. There needs to be a willingness of stakeholders to share data and to invest, if necessary, in data collection if the methodology is to produce the kind of evidence that is strong enough to inform decision making and meaningful advocacy.

The strengths of the methodology are that, together with the conceptual narrative and framework, it facilitates the convening of stakeholders across institutional and organisational boundaries and offers a strong basis for planning and monitoring system change based on a joint examination of assumptions, costs, and outcomes, acknowledging limitations in the available evidence. The potential weaknesses of the methodology are that it is resource intensive, both analytically and organisationally, because it relies on careful collection, analysis, and interpretation of various types of data. The initial pilots in Romania and Malta attest to this. Furthermore, its application may not always offer a clear-cut answer on ‘what is best to do’ – improving the guidance on how to select the interventions being compared, how to value the relevant outcomes, and how interpret the findings remain priorities for future development.

Overall, the methodology does not aim to offer ‘hard’ answers and should not be attempted for such use. Instead, it aims to foster evaluative thinking in the field of child services through structured, system-wide, robust, and context-specific analyses that balance a rights-based foundation and outcome focus with fundamental elements of evidence-informed decision making in the public sector.
Annex A  Examples of approaches to grading evidence

We outline below three approaches for distinguishing between levels of evidence in order to illustrate available approaches to address this issue. A crucial caveat relates to the use of such approaches. As Petticrew and Roberts noted36 (see Section A.3 below), the concept of a ‘hierarchy of evidence’ can be problematic for social interventions, as well as public health interventions, for two reasons:

- first, ‘gold standard’ methodologies do not necessarily and consistently produce different results from ‘second tier’ methodologies; and
- second, the focus should be on the specific question at hand, and some methodologies are by design more apt at answering certain questions than others.

A.1  Oxford Centre for Evidence-Based Medicine (OCEBM)

The OCEBM model presents a hierarchy of the likely best evidence on the benefits of a given medical treatment. Its primary aim is to assist clinicians to conduct their own rapid appraisals, but does not aim to give either definitive answers or recommendations.37 The hierarchy is presented as a step-by-step guide which walks users through a series of questions associated with five levels of evidence, with Level 1 as the strongest and Level 5 as the weakest. The hierarchy matrix is presented below.38

- **Rows** represent a series of steps to follow when searching for likely best evidence. The likely strongest evidence is likely to be found furthest to the left of the Levels, and each column to the right represents likely weaker evidence.
- **Columns** represent the types of questions the clinician is likely to encounter in the order the clinician will encounter them. For example, the first question a clinician might want to ask regards prevalence (how common is it?). Then, they might like to know whether the diagnostic test was accurate. Next, they should wonder what would happen if they did not prescribe a therapy, and whether the likely benefits of the treatment they propose outweigh the likely harm.

Irrespective of the specific questions, be they relating to the value of diagnostic tests or therapies, evidence is the strongest (Level 1) when it is informed by local surveys or systematic reviews of studies. Conversely, it is weakest when it is informed by mechanism-based reasoning or case series.

---

<table>
<thead>
<tr>
<th>Question</th>
<th>Step 1 (Level 1*)</th>
<th>Step 2 (Level 2*)</th>
<th>Step 3 (Level 3*)</th>
<th>Step 4 (Level 4*)</th>
<th>Step 5 (Level 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>How common is the problem?</strong></td>
<td>Local and current random sample surveys (or censuses)</td>
<td>Systematic review of surveys that allow matching to local circumstances**</td>
<td>Local non-random sample**</td>
<td>Case-series**</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>Is this diagnostic or monitoring test accurate?</strong> (Diagnosis)</td>
<td>Systematic review of cross sectional studies with consistently applied reference standards</td>
<td>Individual cross sectional studies with consistently applied reference standards and blinding</td>
<td>Non-consecutive studies, or studies without consistently applied reference standards**</td>
<td>Case-control studies, or poor or non-independent reference standard**</td>
<td>Mechanism-based reasoning</td>
</tr>
<tr>
<td><strong>What will happen if we do not add a therapy?</strong> (Prognosis)</td>
<td>Systematic review of inception cohort studies</td>
<td>Inception cohort studies</td>
<td>Cohort study or control arm of randomized trial*</td>
<td>Case-series or case-control studies, or poor quality prognostic cohort study**</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>Does this intervention help?</strong> (Treatment Benefits)</td>
<td>Systematic review of randomized trials or n-of-1 trials</td>
<td>Randomized trial or observational study with dramatic effect</td>
<td>Non-randomized controlled cohort/follow-up study**</td>
<td>Case-series, case-control studies, or historically controlled studies**</td>
<td>Mechanism-based reasoning</td>
</tr>
<tr>
<td><strong>What are the COMMON harms?</strong> (Treatment Harms)</td>
<td>Systematic review of randomized trials, systematic review of nested case-control studies, n-of-1 trial with the patient you are raising the question about, or observational study with dramatic effect</td>
<td>Individual randomized trial or (exceptionally) observational study with dramatic effect</td>
<td>Non-randomized controlled cohort/follow-up study (post-marketing surveillance) provided there are sufficient numbers to rule out a common harm. (For long-term harms the duration of follow-up must be sufficient.)**</td>
<td>Case-series, case-control or historically controlled studies**</td>
<td>Mechanism-based reasoning</td>
</tr>
<tr>
<td><strong>What are the RARE harms?</strong> (Treatment Harms)</td>
<td>Systematic review of randomized trials or n-of-1 trial</td>
<td>Randomized trial or (exceptionally) observational study with dramatic effect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Is this (early detection) test worthwhile?</strong> (Screening)</td>
<td>Systematic review of randomized trials</td>
<td>Randomized trial</td>
<td>Non-randomized controlled cohort/follow-up study**</td>
<td>Case-series, case-control or historically controlled studies**</td>
<td>Mechanism-based reasoning</td>
</tr>
</tbody>
</table>

* Level may be graded down on the basis of study quality, imprecision, indirectness (study PICO does not match questions PICO), because of inconsistency between studies, or because the absolute effect size is very small; Level may be graded up if there is a large or very large effect size.

** As always, a systematic review is generally better than an individual study.

A.2 California Evidence-Based Clearinghouse for Child Welfare

This organisation has developed a guide for selecting and implementing evidence-based practices for child and family serving systems.\(^{39}\) The guide is accompanied by a typology of evidence which distinguishes between Anecdotal, Descriptive, Pre-Experimental, Quasi-Experimental, and Experimental.

---

**Anecdotal**

Information based heavily or entirely on casual observations or personal testimony rather than rigorous or scientific analysis.

Example: A program participant reports on how parent training has helped him become a better father.

---

**Descriptive**

Data on participant characteristics, the numbers served by the program, case studies, or observational studies.

Example: A program reports on their participant demographics and highlights stories of successful cases.

---

**Pre-Experimental**

Research that does not have a control group (i.e., comparison group). This can include measuring outcomes of a single participant posttreatment or one group of participants pretest/posttest.

Example: At the end of the mentor program, youth report a higher level of confidence compared to their entry scores.

---

**Quasi-Experimental**

Measures outcomes across one or more groups of program participants and a control group without participants randomly assigned to a group. Participants complete standardized assessments prior to entering the program and after completing the program.

Example: The class that completed the program showed significant improvement in behavior compared to other classes in the same grade.

---

**Experimental**

A Randomized Controlled Trial (RCT) randomly assigns participants to either the program or a control group. Any differences seen in the groups at the end can be attributed to the differences in the intervention alone, and not to bias or chance.

Example: Comparisons of both groups at pretest and posttest show that the program group experienced less parenting stress than those in the control group.

---

Petticrew and Roberts\textsuperscript{40} argued for a typology rather than a hierarchy, and developed a flexible matrix-like approach which relates the value of the study design to specific questions and with an application to social interventions in children.

<table>
<thead>
<tr>
<th>Research question</th>
<th>Qualitative research</th>
<th>Survey</th>
<th>Case-control studies</th>
<th>Cohort studies</th>
<th>RCTs</th>
<th>Quasi-experimental studies</th>
<th>Non experimental evaluations</th>
<th>Systematic reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effectiveness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does this work? Does doing this work better than doing that?</td>
<td>+</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+++</td>
</tr>
<tr>
<td><strong>Process of service delivery</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How does it work?</td>
<td>++</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+++</td>
</tr>
<tr>
<td><strong>Salience</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does it matter?</td>
<td>++</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+++</td>
</tr>
<tr>
<td><strong>Safety</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Will it do more good than harm?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td></td>
<td></td>
<td>+++</td>
</tr>
<tr>
<td><strong>Acceptability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Will children/parents be willing to or want to take up the service offered?</td>
<td>+</td>
<td>+</td>
<td></td>
<td>++</td>
<td>+</td>
<td></td>
<td></td>
<td>+++</td>
</tr>
<tr>
<td><strong>Cost effectiveness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is it worth buying this service?</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+++</td>
</tr>
<tr>
<td><strong>Appropriateness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is this the right service for these children?</td>
<td>++</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>++</td>
</tr>
<tr>
<td><strong>Satisfaction with the service</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are users, providers, and other stakeholders satisfied with the service?</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

A.4 Standards of Useful Evidence, Dartington Social Research Unit

A. Evaluation quality

Good enough

A1. One randomised controlled trial (RCT) or two quasi-experimental design (QED) evaluations (initial quasi-experimental evaluation and replication) with the following characteristics (see A1a–A1e):

A1a. Assignment to the intervention is at a level appropriate to the intervention.

A1b. There is use of measurement instruments that are appropriate for the intervention population of focus and desired outcomes.

A1c. Analysis is based on ‘intent to treat’.

A1d. There are appropriate statistical analyses.

A1e. Analyses of baseline differences indicate equivalence between intervention and comparison.

A2. There is a minimum of one long-term follow-up (at least six months following completion of the intervention) on at least one outcome measure indicating whether results are sustained over time.

A3. There is a clear statement of the demographic characteristics of the population with whom the intervention was tested.

A4. There is documentation regarding what participants received in the intervention and counterfactual conditions.

A5. There is no evidence of significant differential attrition.

A6. Outcome measures are not dependent on the unique content of the intervention.

A7. Outcome measures reflect relevant outcomes. Requires evidence that one or more of the outcome measures reflects one or more relevant outcomes.

A8. Outcome measures are not rated solely by the person or people delivering the intervention.

Best

A9. There are two RCTs or one RCT and one QED evaluation (in which analysis and controls rule out plausible threats to internal validity). Requires evidence that at least two RCTs or one RCT and one QED evaluation were conducted on the intervention in question and, critically, that they meet the methodological requirements spelled out in all ‘good enough’ evaluation quality criteria (A1 – A8).

A10. The evaluation results indicate the extent to which fidelity of implementation affects the impact of the intervention.

A11. Dose-response analysis is reported.

A12. Where possible or appropriate there is analysis of the impact on sub-groups (e.g. do the results hold up for different age groups, boys and girls, ethnic minority groups?)
A13. There is verification of the theoretical rationale underpinning the intervention, provided by mediator analysis showing that effects are taking place for the reasons expected.

B. Impact

Good enough

B1. There is a positive impact on a relevant outcome. Requires evidence that in a majority of studies complying with the ‘good enough’ evaluation quality criteria set out in section A, programme group participants did better relative to the control group participants on a relative outcome, and that the difference is statistically significant.

B2. There is a positive and statistically significant effect size, with analysis done at the level of assignment (or, if not, with appropriate correction made.)

   or

   There is a reported sample size of 0.2 with a sample size of more than 500 individuals across all studies.

B3. There is an absence of iatrogenic effects for intervention participants. (This includes all sub-groups and important outcomes.)

Best

B4. If two or more RCTs or at least one RCT and one QED evaluation have been conducted, and they meet the methodological criteria stipulated in section A (see criterion A9), there is evidence of a positive effect (criterion B1) and an absence of iatrogenic effects (criterion B3) from a majority of the studies.

B5. There is evidence of a positive dose-response relationship that meets the methodological standard stated in A11.

C. Intervention specificity

Good enough

C1. The intended population of focus is clearly defined.

C2. Outcomes of intervention are clearly specified and meet one of the relevant outcomes.

C3. The risk and promotive factors that the intervention seeks to change are identified, using the intervention’s logic model or theory explaining why the intervention may lead to better outcomes.

C4. There is documentation about what the intervention comprises.

Best

C5. There is a research base summarising the prior empirical evidence to support the casual mechanisms (risk and protective factors) that underlie the change in outcomes being sought.
D. System readiness

Good enough

D1. There are explicit processes for ensuring that the intervention gets to the right people.

D2. There are training materials and implementation procedures.

D3. There is a manual(s) detailing the intervention.

D4. There is reported information on the financial resources required to deliver the intervention.

D5. There is reported information on the human resources required to deliver the intervention.

D6. The intervention that was evaluated is still available.

Best

D7. The intervention is currently being widely disseminated.

D8. The intervention has been tested in ‘real world’ conditions.

D9. Technical support is available to help implement the intervention in new settings.

D10. Absolute financial investment is stated.

D11. There is a fidelity protocol or assessment checklist to accompany the intervention.

Note: More detailed explanations of these standards exist for those completing programme reviews: they are available from Nick Axford at the Social Research Unit, Dartington (naxford@dartington.org.uk).