



ARTICLE

DOI: 10.1057/s41599-018-0139-z

OPEN

Developing theories of change for social programmes: co-producing evidence-supported quality improvement

Deborah Gbate  ¹

ABSTRACT For much of the past two decades, expensive and often imported evidence-based programmes (EBPs) developed by clinician-researchers have been much in vogue in the family and parenting support field, as in many other areas of social provision. With their elaborate infrastructures, voluminous research bases and strict licensing criteria, they have seemed to offer certainty of success over less packaged, less well-evidenced locally developed approaches. Yet recently, evaluation research is showing that success is not assured. EBPs can and regularly do fail, at substantial cost to the public purse. In times of severe resource pressure, a pressing question is, therefore, whether lower cost, home-grown, practitioner-developed programmes—the sort often overlooked by policy-makers —can deliver socially significant and scientifically convincing outcomes at lower cost and at least on a par with their better resourced cousins. This paper shows how the application of techniques increasingly used in implementation science (the science of effective delivery) could help level the playing field. Processes for doing this including co-produced theory of change development and validation are illustrated with reference to the *Family Links Ten Week Nurturing Programme (FLNP-10)*, a popular manualised group-based parenting support programme, designed and disseminated since the 1990s by a UK-based purveyor organisation. The paper draws out general principles for formulating and structuring strong theories of change for practice improvement projects. The work shows that novel application of implementation science-informed techniques can help home-grown programmes to compete scientifically by strengthening their design and delivery, and preparing the ground for better and fairer evaluation.

¹The Colebrooke Centre for Evidence and Implementation, London, United Kingdom. Correspondence and requests for materials should be addressed to D.G. (email: dghate@cevi.org.uk)

Introduction: evidence-based and home-grown parenting support in policy and practice context

For at least the past two decades, commissioning of social programmes across many sectors and countries has been dominated by pursuit of the goal—a kind of holy grail—of ‘evidence-based’ interventions: a ‘scientific solution’ to a wide range of social ills that will certainly work, even in the face of intractable problems, and almost irrespective of context, if only they are designed and delivered according to certain key principles (Dick et al., 2016). In no field has this been truer than in parenting support and education, and during the past 20 years or so, a regrettable divide has opened up in the landscape of parenting programme provision.

On one side of the divide stand what we might call indigenous or ‘home-grown’ programmes; the poor, unfashionable relations. These are generally practitioner-developed interventions (Kaminsky et al., 2008), often designed by those working in the field as parent support professionals. They tend to have certain characteristics: they are frequently popular with practitioners and local users; they are often well-integrated into wider local systems of usual care; they are relatively small scale, not having undergone extensive national scale-up; have minimal accreditation requirements; are delivered by a staff from mixed backgrounds with varying levels of professional training; are relatively low-cost; and are often supported only partially by evidence of effectiveness. They also often struggle to gather the substantial resources required to rectify this absence of evidence.

On the other side of the divide stand the so-called evidence-based programmes or interventions (EBPs, EBIs), often developed in the United States by clinician-researchers with scientific backgrounds. These have often already been scaled up nationally and internationally. They tend to have well-developed implementation infrastructures including specific licensing and accreditation requirements; detailed manuals and delivery guidance; are relatively high-cost to deliver; and are usually supported by a substantial scientific evidence base, albeit often gathered in distant settings with different systems of service provision. They tend to stand alone, presented as distinct ‘pockets of excellence’ within the wider landscape of services as usual.

In the UK, influential policy makers in successive administrations have been enthused by the reported results from overseas EBPs, and over time, a series of such programmes have been introduced including particularly from the USA and Australia, at considerable investment of local effort and public money. EBI developers and others have held out the promise of strong returns on investment and programmes have often been ‘sold’ to policy makers and commissioners as producing assured success if implemented as designed. Such has been the level of confidence that financial investment instruments have been developed associated with the programmes, for example, social impact bonds attached to Multi-Systemic Therapy (Edmiston and Nicholls, 2018), which promise to pay social investors financial returns for pre-specified results. A veritable commissioning ‘bandwagon’ (Dick et al., 2016) has resulted, but—as it is slowly becoming clear—often with mixed or outright disappointing results in replication in new settings (e.g., Fonagy et al., 2018; Cottrell et al., 2018; Forrester et al., 2018; Humayun et al., 2017; Robling et al., 2015; Little et al., 2012; Biehal et al., 2010).

Of course, this is not the whole picture: there has also been policy-driven investment of substantial public money in a number of home-grown programmes designed in the UK that cannot be described as fully (or sometimes even partially) evidence-based, also with mixed (Melhuish et al., 2008) or disputed results (Crossley, 2015). Conversely there have been consistently successful attempts to introduce some EBPs in some arenas (e.g., Morpeth et al., 2017). But the policy dilemma is thus: off-the-shelf EBPs are expensive, with sometimes onerous conditions attached

to their implementation. They are often hard to implement in new jurisdictions ‘as designed’ and the process of adapting them, if permitted, is slow and requires fresh investment in research to meet the required standards of evidence. When they fail to improve on the results of services ‘as usual’, as they not infrequently do, the costs are large and the consequences attract negative opinion. On the other hand, while local practitioner-developed programmes are certainly less expensive to deliver, the lack of convincing evidence of effectiveness has been—especially during the years of rhetoric regarding ‘evidence-based policy-making’ (Hammersley, 2005) and ‘outcomes-based commissioning’—off-putting to central government in particular, as well as to the scientific community. Yet resources for children’s social care are diminishing and demand is rising (All Party Parliamentary Group, 2017) and it seems possible that the wave of enthusiasm for commissioning high-cost EBPs may have crested for the time being.

Given this ‘crisis of replication’ (Baker, 2016; Grant and Hood, 2017) for imported EBPs, we need to look again at what we can achieve with limited resources, and how we can improve the services that already exist on our own doorstep. A key question, therefore, is: can promising low-cost indigenous programmes be supported to improve their effectiveness and their associated evidence base, so that even when resources are most stretched, high-quality support can still be made available to families in vulnerable communities? In this article, it is suggested that constructs and methods being championed by the fields of implementation and improvement science can successfully be deployed to do just that, and one approach to undertake quality improvement collaboratively is described.

Background to the FLNP-10

The *Family Links 10-Week Nurturing Programme (FLNP-10)*¹ is a 10-week community-based, ‘purveyor-supported’² parenting support programme, designed to be delivered by trained parent group leaders who are employed by or contracted to provider agencies. The programme is popular, respected and widely used across local authorities in England, Wales and Northern Ireland. Provider agencies are predominantly local authorities, with some voluntary organisations and independent consultancies also trained by Family Links to deliver the programme. 10-Week Nurturing Programme courses are usually delivered within community-based venues such as children’s centres and schools.

Originating in a US-developed intervention known as the *Nurturing Parenting Programmes*³ (see Bavolek, 2000), the FLNP-10 has been adapted and developed independently over many years by practitioners at Family Links for delivery in the UK. It now ranks, therefore, as an indigenous programme rather than an import from the canon of international off-the-shelf EBPs.

Evidence status of the FLNP-10

The UK programme has been the subject of a number of research studies of varying size, scope and methodologies including (unusually for an indigenous programme) a large-scale randomised control trial (RCT) funded by government (Simkiss et al., 2013). Results have been mixed, following a pattern familiar in much UK parenting support research (Moran and Gbate, 2013) of positive results from qualitative research but inconclusive results in larger scale quantitative research. Local and generally small scale studies have consistently yielded results showing positive changes for programme beneficiaries. However, the independent RCT produced a less positive picture. While finding a consistent trend for more positive outcomes for parents and children in the treatment arm, it also detected no results that

reached the levels required for statistical significance, although the authors noted that methodological issues may have affected the results.

Perhaps Campbell's Law was at work here⁴; or of course, it may be the case that the programme is ineffective. The research does not confirm either way at this point. But given the warm regard in which the programme is held as well as copious non-experimental data pointing to effectiveness, when faced with the disappointing and unexpected results of the RCT the purveyors queried whether methodological weaknesses in the trial design, and specifically in the make-up of the sample used for the treatment arm, might have reduced the size of differences between the two groups (Mountford and Darton, 2013). An additional or alternative explanation—one that reflects growing insights from the field of implementation science—leads one to question the selection of the specific outcome measures used in the trial, which borrowed their design from a contemporary, much larger trial of the well-known government-funded *Sure Start* programme in the UK. In particular, it appears reasonable to question whether these outcomes most appropriately reflected the programme's actual implementation model and objectives.

This was especially so given that there was, at the time of the trial's design, nothing in writing that set out the theory of change employed by the programme, and nothing that traced the expected logic of the pathway to change, from programme inputs and activities through implementation processes to eventual outcomes for families. This would have made it difficult to design the suite of outcome measures, no matter what the style of evaluation. Certainly the FLNP-10 is nothing like the *Sure Start* programme (except in that it is also a parenting support programme), and measures to capture change at the child level seemed especially questionable, given that unlike many components of *Sure Start* the implementation model of the FLNP-10 programme does not involve any direct contact or intervention with children themselves, but works solely through parents.

The decision to develop a theory of change: insights from implementation science

Notwithstanding these reservations, this article is not intended as a critique of the specific RCT in question, nor of RCT methods in general. There is already a growing literature on the challenges of the latter (e.g., Cartwright, 2011; Cartwright and Hardie, 2012; Pearce et al., 2015), and the authors of the Nurturing Programme RCT themselves have since written on the difficulties and inadequacies of measurement in this field (Stewart-Brown et al., 2011). In the broad inter-disciplinary community now working globally in implementation or improvement sciences, where the focus is on the systematic study and testing of implementation processes and innovations in order to expand the knowledge base about what works (UK Implementation Society, 2017), it is in any case becoming widely accepted that experimental methods used alone may lack the sophistication of perspective to be effective tools for assessing the results of interventions that take place in complex adaptive or responsive systems (Ghate, 2015), as all community-based health and social interventions do (Ghate, 2015; Kainz and Metz, 2016; Chambers et al., 2013; Hawe et al., 2004; Mowles, 2014).

Rather, the foregoing comments are intended to offer an all-too-typical real world example of the impossibility of fair independent evaluation of interventions that have no articulated causal model. This illustrates the general problems of evaluation for programmes and services whose theory of change is implicit, and contained in the 'tacit knowledge' of its developers and purveyors rather than rendered explicit through external articulation. Without a clear roadmap of precisely how the programme

designers intend to arrive at their goals, external evaluators are working in the dark: they must make assumptions and join dots, most especially as they try to intuit how the implementation model—the engine that drives an intervention—works in practice. Inevitably, given the speed at which most real world research projects must proceed, misapprehensions may result.

To the initiated, these comments may seem self-evident. It is important to note, however, that Family Links were not at all unique in having no clearly articulated theoretical model of their work: in fact, among practitioner-led social programmes developed in the field rather than designed in the laboratory, this is very common. The absence of an articulated theory of change is also not only a problem of social interventions; health interventions also suffer from this (Davis et al., 2015). For example, in the field of health behaviour change, a recent meta-analytic study found that only just over one fifth of published studies explicitly drew on theories of change in their analysis (Davies et al., 2010). This is in spite of widely-cited work, over many years, flagging the importance of theorising change prior to summative evaluation, including by those working in so-called Realist Evaluation (Pawson and Tilley, 1997; Pawson, 2013). The imported EBPs beloved of UK policy entrepreneurs over the past three decades almost always had an articulated theory of change, and we have previously argued that this was not at all co-incidental to their ability to establish convincing evidence of efficacy (Utting et al., 2007). Yet it is only relatively recently that there has been widespread interest among commissioners of social interventions in having designers and purveyors set out a theory of change as an *indispensable* element of effective delivery. As implementation and improvement scientists have increasingly understood, a causal model may be constructed from any one of a number of directions (Kainz and Metz, 2016), but constructed it must be.

Recognising this, it was agreed that the lack of clearly articulated theory of change for the FLNP-10 needed to be addressed. The remainder of this article explores the wider conceptual basis for this work. It then sets out the method used and describes a process that could, in principle, be used by any social programme wishing to strengthen its implementation quality and its evaluability. This work was successfully undertaken post hoc (working with the programme 'as is') rather than a priori (working with a programme 'to be' while it is being designed), and, therefore, shows that even existing and well-established programmes can undertake improvement work using these methods, even if that was not their historical starting point.

Theories of change: a quality improvement tool

Components, principles and variants. A theory of change is in essence no more than a planned route to outcomes: it describes the logic, principles and assumptions that connect what an intervention, service or programme does, and why and how it does it, with its intended results. Some authors refer to this as 'programme theory' (e.g., Rogers, 2008). It is a formal and explicit articulation of the assumptions that underpin the rationale and design of a programme or intervention, and explains why it is reasonable to expect the programme to achieve change for service users (although as a 'theory', it does not, of course, offer any guarantees of effectiveness, only hypotheses). In intervention or implementation science, it is helpful to conceptualise this as a 'pathway to change' along which service users should travel, moving from their initial presenting needs or problems to the final positive outcomes that the programme hopes to achieve (Ghate, 2016; Renger and Hurley, 2006; Weiss, 1997; Rogers, 2008; Hawe, 2015). A *logic model* (a term sometimes confusingly used interchangeably with 'theory of change') is a pictorial representation of the theory (Hawe, 2015), usually in short form.

It may be produced as part of, or as an output of, the process of development of the theory of change, as the ‘summarised theory of how the intervention works, (usually in diagrammatic form)’ (Rogers, 2008 : 30).

What components are key to a plausible theory of change?

Theories of change as described in what has now become a copious literature can take different forms. Ideally, in addition to the basic elements of *needs* (the initial problem being addressed), *inputs* (resources), *outputs* (intended activities) and *outcomes* (desired changes for service users), a theory of change should also include specification of *implementation outcomes* at practice, organisation or system levels (Ghate, 2015) as well as the *mechanisms of change*, (Weiss, 1997) both of which are required to produce the intended *final outcomes* for ultimate beneficiaries.

Needs may be framed as ‘problems’ or deficits that the intervention addresses, but they can also take the form of positive strengths (in the case of interventions aiming to boost resilience, for example). Antecedent or *root causes* of the specific problem or need being targeted are also helpful to include in the theory, even if they are systemic and distal in nature and causally ‘out of range’ for the intervention in question to address (e.g., poverty, social inequality). They are important for promoting recognition of the multi-factorial nature of social or health needs. Root causes are often ‘wicked problems’ (ones that resist straightforward or single-factor solutions, (Rittel and Webber, 1973), or historical features of service users’ situation that cannot be eliminated even by the most sophisticated service response, (e.g., past trauma or abuse). But including them in a theory of change encourages providers and commissioners to recognise the intervention in its systemic context and to see it as one part of wider picture of impetus for change, rather than as an isolated change-agent. It prevents our unwittingly under-stating the complexity of the problem (Renger and Hurley, 2006) and it may help to keep expectations realistic by resisting the unhelpful impression that a single intervention or programme (no matter how impressive or elaborate) can be a ‘silver bullet’ for complex, multiply determined social problems.

Resources and *activities* should carefully be specified; these form the heart of the implementation model but are often poorly described, especially if the programme or intervention concerned is part of or relies on elements of established practice, delivered by established organisations. Specifying this ‘operational form’ of an intervention is critical to accurate visioning of outcomes. *Activities* refers to the content (or curriculum or rubric) of the intervention, for example, specific messages to be transmitted, specific practices to be taught, specific methods to be used, and ordering of sessions. *Resources* refers to the infrastructure necessary to produce that content (type, number and competencies of staff or volunteers; training required to produce and support them; equipment and facilities; ancillary services such as transportation, crèches etc). ‘Dosage’ (how many sessions, how long for, over what period) can also be specified here, in order to set out what level of exposure to the intervention is thought to be necessary to produce the required results for the average user.

Implementation outcomes are often overlooked during the specification of how a programme is intended to work, but unless an intervention involves direct delivery to users without the intermediation of staff or volunteers—for example, provision of simple equipment, say, delivering recycling collection bins to domestic households—implementation outcomes are almost always required in order to produce final ‘treatment outcomes’ for service users. Implementation outcomes—like all outcomes—should be conceptualised as changes arising as a result of the intervention (Moran and Ghate, 2013), in this case concerned

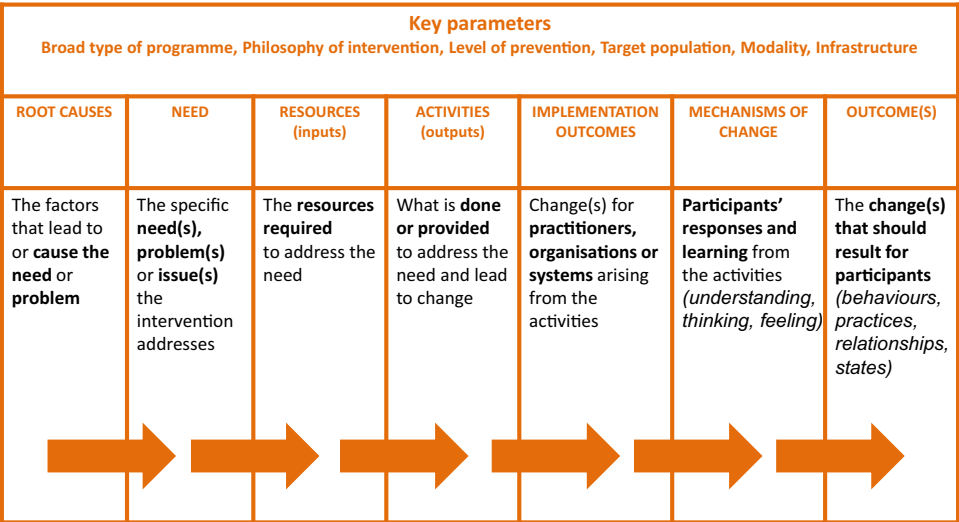
with the changed professional behaviours required from individual practitioners, from teams and organisations, and from the wider systems in which these are embedded, which demonstrate that new practices are being properly and effectively employed and supported (see Ghate, 2015, for more on implementation outcomes)⁵. These might include changes in the practices or behaviours of individual staff when working with service users; changes in how organisations arrange themselves to support the delivery of an intervention; or changes at a higher systems-level—for example developments in cross-agency collaboration, or supportive policy or commissioning changes.

Mechanisms of change, defined as learning and value shifts for ultimate beneficiaries are also frequently left unspecified in theories of change. However, they too are critical to the specification of causality (Davies et al., 2010), where it is essential to consider ‘the mechanisms that intervene between the delivery of the programme service and the occurrence of the outcomes of interest’ (Weiss, 1997, p 46) in order to fully model and understand how a programme is intended to work. These are usually learning changes of a cognitive, attitudinal, emotional or mental variety. They can be thought of as mediating (facilitating, enabling, re-inforcing) the ultimate outcomes. They are the essential, intermediate step along the logical pathway to change, which in most social or health interventions culminates in targeted *treatment or service outcomes* in the form of changes in service users’ behaviours or changes in their state.

Lastly, overarching everything else, specification of the key *parameters* is also vital within a theory of change: these are defined as the boundaries within which the intervention is intended to function (for example, who it is for, in what settings and at what level of prevention). This prevents the intervention being delivered inappropriately, and allows robust testing of how it works when the parameters are manipulated, including during replication and scaling. Sometimes the parameters will reflect the logic of the invention perfectly; but sometimes they are also shaped by political or economic realities or ideological positions—for example, where services potentially useful to many may be offered only to a sub-group of particular recipients, perhaps reflecting cost and resource constraints, or views about who is deserving of help. Parameters can also reflect the macro-theoretical level—for example, the broad type of intervention (e.g., behavioural, cognitive, technological, pharmacological etc), or an overarching philosophy of care or of human behaviour (attachment-based, restorative, therapeutic etc) underpinning the intervention design.

In setting out the elements of the theory, certain overarching principles emerge as important for robustness and utility. One is that the specification should be as clear and as concrete as possible (Weiss, 1995). And perhaps the most important prescription for drawing up a theory of change, including a summary logic model, is that it should be as simple as possible. The guiding principle should be *parsimony* (Davies et al., 2010): ‘minimum critical specification, defining the fewest programme elements possible to produce the desired effect’. (Bradach, 2003, p 21). According to Bradach, one way to define core operational elements of a theory of change (see below) is ‘to ask whether varying an element would diminish the value a programme creates’.

What form should a theory of change take? Whetten notes that theory should be built around two elements, *description* and *explanation* (Whetten, 1989). Description (the form and context of the thing being theorised) includes antecedent causes of the need addressed by a programme, and consideration of the context in which the need arises, as well as a clear description of the



© 2017, The Colebrooke Centre for Evidence and Implementation

Fig. 1 Schematic theory of change model. This Figure is covered by the Creative Commons Attribution 4.0 International License. Reproduced with permission of The Colebrooke Centre for Evidence and Implementation; © Colebrooke Centre for Evidence and Implementation, all rights reserved

intervention proposed, who it is intended for, and how it should be delivered in practice, and what results should be expected (as Whetten puts it: the ‘material ‘in the boxes’ of a flow diagram’). Explanation (the reason why the desired change is expected to occur) is implicit in the ‘arrows that connect the boxes’, and in elucidation of the mechanisms of change. Other definitions also stress the significance within theory specification of ‘how phenomena relate to one another’ (Davis et al., 2015); and the influential Realist model originally developed to help shape better evaluation emphasises clarity of connections between three key elements: context, mechanisms and outcomes (C-M-O), where certain contexts are considered to ‘trigger’ mechanisms that in turn, generate outcomes (Wong et al., 2012).

When constructing the theory, careful thought should be given to the plausibility of the arrows, since this ‘causal pathway’ is what will produce the results, and what will be tested by subsequent evaluation. Ideally, there should be pre-existing evidence from other change initiatives to support the hypothesis that the ‘boxes’ can indeed be connected logically in the way set out by the theory; in other words, that they have predictive significance (Davies et al., 2010).

Should the form be linear, or cyclical? Should it be open, or closed? Theories of *behaviour* tend to be visualised in a linear form (Davis et al., 2015)—and it will be noted that the schematic in Fig. 1 has a distinct linearity. Perhaps because of the influence of this paradigm, the alluring certainty of propositional logic (Mowles, 2014) and because many social interventions do aspire to influence behaviour for participants, numerous ‘text book’ examples of how to draw up a theory of change have adopted this form⁶. Indeed, any glance at the hundreds of examples of images for theory of change models produced by a search of the Web using any browser will illustrate the dominance of this form: almost all models of social programmes assume linear progression of one kind or another.

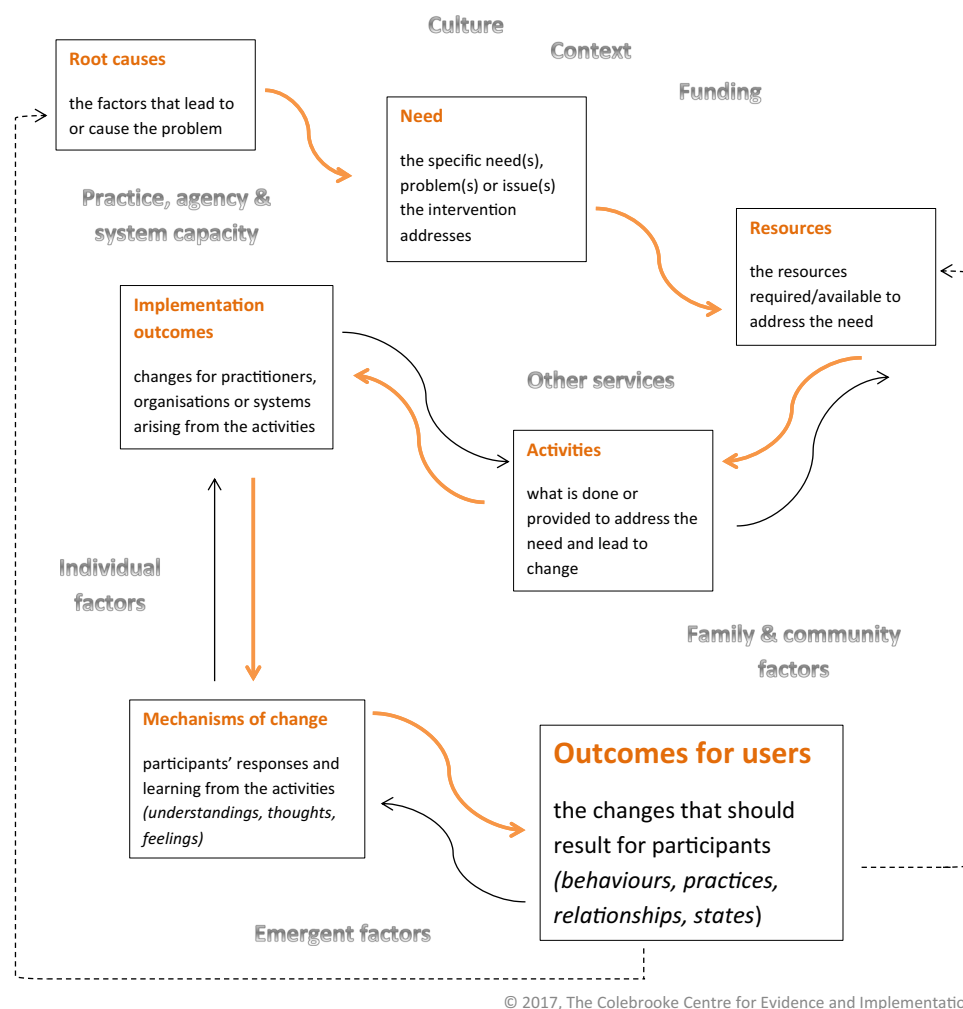
Linearity has the advantage of relative simplicity. When starting out on what can feel like a daunting intellectual process, it may help the task of ordering the thinking regarding different components of the model and how they inter-connect. A linear logic model is an improvement on no logic model at all, and a linear schematic is (probably) a sensible place to start. But setting out the theoretical model as sequence of inter-linked dimensions should not mislead us into thinking that the process of

implementation will proceed in a similarly neat linear form, reliably propelling the programme participants towards the desired results. The evidence tells us that this is not how implementation works (Fixsen et al., 2005).

As we increasingly recognise, theories of *change* more broadly construed need to reflect the real world, and must accommodate for the concept of ‘emergence’: something ‘that arises from but is not reducible to the sum of all activity that comprises interactive agents acting locally’ (Mowles, 2014, p 166)—in other words, the unpredictable results of human beings interacting with one another and with the structures and systems that surround them. As Mowles (2014) has persuasively written in the context of evaluation methods but with striking applicability to implementation theory and practice, even when interventions seem ‘simple’ or straightforward, their operation in the real world is almost certainly very far from this. Given the state of non-equilibrium and non-linearity that characterises all social and human systems, it is probably time we threw out the closed, linear format for the purposes of implementation and improvement planning, and wrestled more actively with the complex responsive and contingent processes that characterise actual implementation in actual service settings (Braithwaite et al, 2018); see Fig. 2.

Why is a theory of change useful and important? There is now widespread international consensus regarding the scientific importance of elucidating the theory of intervention, for example, as noted in recent UK Medical Research Council guidance on process evaluation, ‘a strong understanding of the theory of intervention is a pre-requisite for a meaningful assessment of whether the (delivered) intervention remained consistent with its underlying theory’ (Moore et al., 2015, p 41). With the roadmap thus created, the assumptions on which the programme is based can be tested and evaluated more fairly. Each key element can (in principle) be measured to check that it was present in delivery, and expected outcomes can be appropriately measured, for example, before, during and after the service has been taken up.

Increasing ‘evaluability’ is one clear benefit of having an articulated theory of change, but there are others, set out in the Box 1. Most important of all is the use to which the theory can be put for programme improvement purposes. Once it is rendered explicit, the programme theory can be tested, and then refined or amended to take account of the results of testing. These kinds of



© 2017, The Colebrooke Centre for Evidence and Implementation

Fig. 2 Real World, non-linear theory of change model. This Figure is covered by the Creative Commons Attribution 4.0 International License. Reproduced with permission of The Colebrooke Centre for Evidence and Implementation; © Colebrooke Centre for Evidence and Implementation, all rights reserved

iterative processes should result, over time, in more effective and more cost-efficient programmes for families. Another of the most important benefits of having a theory of change—its importance for replicability—is especially relevant for programmes wanting to scale-up and expand their reach, since: ‘Without a strong theory of change, replication becomes extremely difficult because it is impossible to determine what is working or why’ (Bradach, 2003, p 21; see also Baker, 2010)

As Haynes and colleagues (2016) note, assessing theoretical fidelity also helps the intervention to be further developed and potentially streamlined as experience accumulates through using it in different contexts. This latter benefit is a considerable asset to the scaling process over time. It also helps in the ‘here and now’ of programme implementation for example in particular by illuminating *activity traps* in which staff may have unwittingly become involved—defined as inefficient or ineffective practices which ‘occur when the activities of the intervention target the symptoms of the problem rather than the conditions leading to the problem’ (Renger and Hurley, 2006; see also Renger and Titcomb, 2002). Most well-established programmes can probably find examples of these in their own operations:

practices or processes that are faithfully transacted because this is ‘how we do it’, without a clear analysis of whether they are material to the effectiveness of the intervention.

The opposite of activity traps are *core components* (elements of content, or of implementation, or both) that are believed to be

essential to the programme’s effectiveness (Blase and Fixsen, 2013; Haynes et al., 2016). Again, the process of articulating the theory of change is helpful for identifying what are hypothesised to be the ‘non negotiables’—whether defined by form or function (Hawe et al., 2004)—ensuring that if local adaptations are made, they are sensitive and disciplined and do not undermine the programme and make it less successful, helping to prevent ‘programme drift’ (slippage or unnecessary modification by providers and practitioners; Chambers et al., 2013)

Limitations and criticisms of theories of change. Although the benefits are clear, a theory of change does not ‘proof’ an intervention against failure. There are certainly pitfalls, and not all commentators are equally positive about theories of change; for example, Mulgan (2016) claims that most theories of change are not in fact ‘theories’ and that they do not necessarily support systematic thinking (although, one might add, they may certainly give the *impression* of systematic thinking). Certainly, a main danger lies in the attempt to simplify what is likely to be a complex reality. Hawe (2015, p 312) notes that the distillation of theory, important though this often is, is problematic: ‘Logic modelling for simple, linear interventions is different from models that attempt to incorporate complexity. This is important because a simple model applied to a complex situation risks overstating the causal contribution of the intervention’.

Box 1: | Benefits of developing a theory of change

1. The **evaluability** of the programme—both for implementation and outcomes—is facilitated, by signposting appropriate metrics
2. The original **intentions** of the programme developers are clearly set out, and are explicit and open to critique
3. The **underlying logic** of the assumptions made in the theory, for example, that undertaking a certain activity will lead to a particular outcome, can be scrutinised
4. The **realism** of the assumptions made by the programme developers can be checked against wider evidence of ‘what works’, to assess the likelihood of the programme being successful
5. Commissioners can check the programme meets their needs; and providers and practitioners delivering the programme can check their own assumptions and the **alignment of their expectations** against the original intentions of the programme developers
6. The **key parameters or boundaries** (e.g., who is the programme for, and under what specific circumstances) can be set out, reducing the likelihood that the programme is used inappropriately or ineffectively
7. **Core components** (of content, or of implementation, or both) that are believed to be *essential* to the programme’s effectiveness can be identified
8. **Activity traps** can be identified and avoided
9. The most important features of the implementation model of the programme can be captured, enabling delivery that adheres to the original model and helping to prevent **programme drift** during maturation and scaling

Mowles (2014) goes further, claiming that there are in fact no such things as simple programmes (nor even, ‘complicated’ ones), but only complex ones. Rogers (2008) observes that logic models (in particular) may instil a false sense of confidence because real world social interventions tend to be complex and multi-level (or at least, taking place within complex systems), whereas logic models are intentionally simplified, compressing rather than fully representing complexity. Logic models in particular should, therefore, be regarded as schematics rather than formulae. It may also be that in scaled up interventions (where multiple sites implement a particular intervention), while the broad theory of change remains the same irrespective of delivery location or setting, somewhat different summary logic models may need to be specified to capture the variation in the local delivery conditions.

Applying a theory of change to the FLNP-10: towards an evidence-supported design

Co-production and iteration. In this section, we describe the process of developing a theory of change specifically for the FLNP-10, by way of illustration of how this process can work in a real world setting. Overall, a key principle was that the process should be co-produced. Otherwise it seemed unlikely that the organisation would feel sufficient confidence in the product and ownership of it for the theory to become useful in their daily practice, or to be regarded as a credible guide for future research and evaluation. Co-production in this case meant that the process should deeply involve key organisational stakeholders (Pfitzer et al., 2013) as well as external contributors, so that practice-based expertise (Durose et al., 2017) could drive the outcomes. The stakeholders in the programme, therefore, worked alongside an implementation specialist and researcher who could help shape and give momentum to the process and in particular provide critical challenge based in knowledge of the wider evidence base on effective parenting support. The process was necessarily iterative and recursive, and had two main stages; stage one formed part of a much larger project to review various aspects of the programme’s functioning and evidence base and resulted in a ‘first cut’ at theory of change; stage two focused on refining and validating the theory of change against wider evidence of ‘what works’, involving some of the same and some new stakeholders. Both stages involved a degree of challenge, especially for the programme stakeholders as we together repeatedly asked the questions ‘why’ and ‘how’ (Renger and Titcomb, 2002) and progressively interrogated long-established and cherished assumptions underpinning how the programme was thought to

work. In the process we together arrived at a refreshed understanding of the purpose of the programme and plausibility of its aspirations. This kind of realignment of thinking and values is now emerging as an important feature of co-created processes (Schneider et al., 2017; Huberman, 1999), which may not only generate new knowledge but may involve transforming present understandings (potentially for all parties) in significant ways. Other authors have noted that it is important, however, that this process of challenge and questioning be conducted ‘in a sensitive and collaborative way, so that it strengthens the programme without dampening enthusiasm or diminishing gut-level commitment’ (Kaplan and Garrett, 2005, p 171).

Identifying content experts. ‘Content experts’ are the people who know the programme from the inside (Renger and Hurley, 2006). For this programme, they included the founding CEO, who had been responsible for bringing the programme to the UK and re-developing it for use in community-based parenting support settings, including co-authoring the manuals used by parents and by group leaders; a new CEO who joined part-way through the process; several senior and longstanding members of staff with responsibility for different aspects of the current programme delivery (quality assurance, liaison with commissioners and providers, resources and support, in-house research); several highly experienced ‘front line’ parent group leaders; and trainers of other group leaders. For timing and resource reasons, we decided not to include service users or participants directly as prior projects on which this project drew in Stage 1 had collated views of service participants over many years. However, service participants are also content experts, so this could have added a further dimension to the work and might have deepened thinking about the implementation model. It remains potential work for the future.

Although knowing ‘from the inside’ makes for deep familiarity and confidence on the part of these informants, for such people, it can also mean that some aspects of the programme’s historical design are simply taken for granted: ‘the way we have always done it’. Critical but friendly challenge is also, therefore, a vitally important part of the process, to surface and explore assumptions. In order to be able to ask searching but sensitive questions it was important for the external consultant to have a good grasp of the programme’s content before the workshops took place, and so she also took part in training for potential parent group leaders to deliver the programme, talked to a number of commissioners and providers and academics familiar with the programme, and

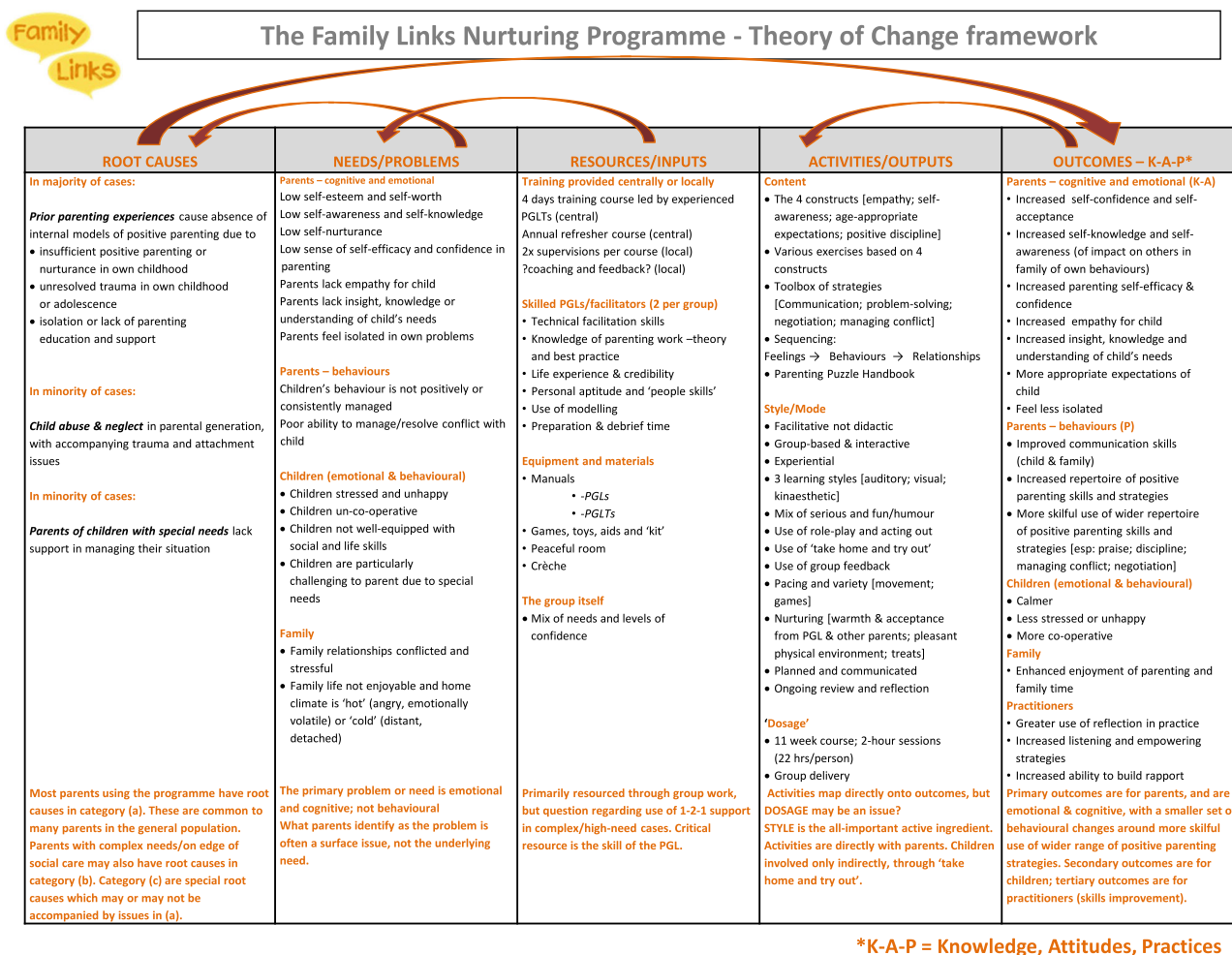


Fig. 3 Stage 1 Draft synthesis model for FLNP-10. This Figure is covered by the Creative Commons Attribution 4.0 International License. Reproduced with permission of The Colebrooke Centre for Evidence and Implementation and Family Links

thoroughly scrutinised the programme manuals and handbooks as part of background preparation.

Stage one: working with the content experts. The external consultant worked alongside programme content experts first, to elucidate thinking and build agreement about how to describe the programme and present its key features, and second, to interrogate underlying assumptions about the casual model and develop the explanation regarding why the programme should be expected to achieve change for parents.

Four separate facilitated sessions in workshop style, each lasting several hours, were carried out with different groupings of content experts. Each of the elements set out in Fig. 1 was sequentially and recursively discussed, gradually populating and refining the model. One of the results of 'challenge' arising from this part of the process was, for example, that the programme's historic framing as (in part) a child abuse prevention intervention began to be called into question. The co-production group agreed that this was no longer substantiated by the actual needs and problems addressed by the contemporary programme, or by the actual activities that were part of the contemporary implementation model. This framing was, therefore, de-emphasised, (and eventually was removed altogether in Stage Two as part of the validation process). This illustrates the value of challenging tacit assumptions that have become buried in the conventional wisdom of the delivery agencies, but which may now be outdated.

The conversations in this stage resulted in an agreed and finalised synthesis model created by the facilitator (see Fig. 3) to represent the different discussions, and a high level logic model summarising Fig. 3 (see Fig. 4).

Stage Two: refining and validating the model. Stage Two, conducted a year later, reviewed and refined the initial model and validated the model by checking its assumptions against the international evidence base on 'what works' in parenting support through an evidence review process. Further workshops were also carried out at this time. One key addition to the first stage visualisation was to populate a column/dimension outlining the hypothesised mechanisms of change, which were agreed to sit logically between two set of activities that are central to the programme, (1) communicating 'four constructs' (principles of optimal parenting on which the programme was based) and (2) teaching and modelling 'four strategies' (ways of putting the constructs into practice in daily parenting). Implementation outcomes were (regrettably) not included in the model. This was partly because at that time, this was not an area that had been thought about at all before in relation to the programme. It was also not an area on which the organisation already had data, and resources were not available to remedy this.

Validation (which we termed 'evidence-mapping') was a key part of the Stage Two process. It was considered necessary to ensure that the final theory of change would be plausibly aligned with what is known about what is achievable through parenting

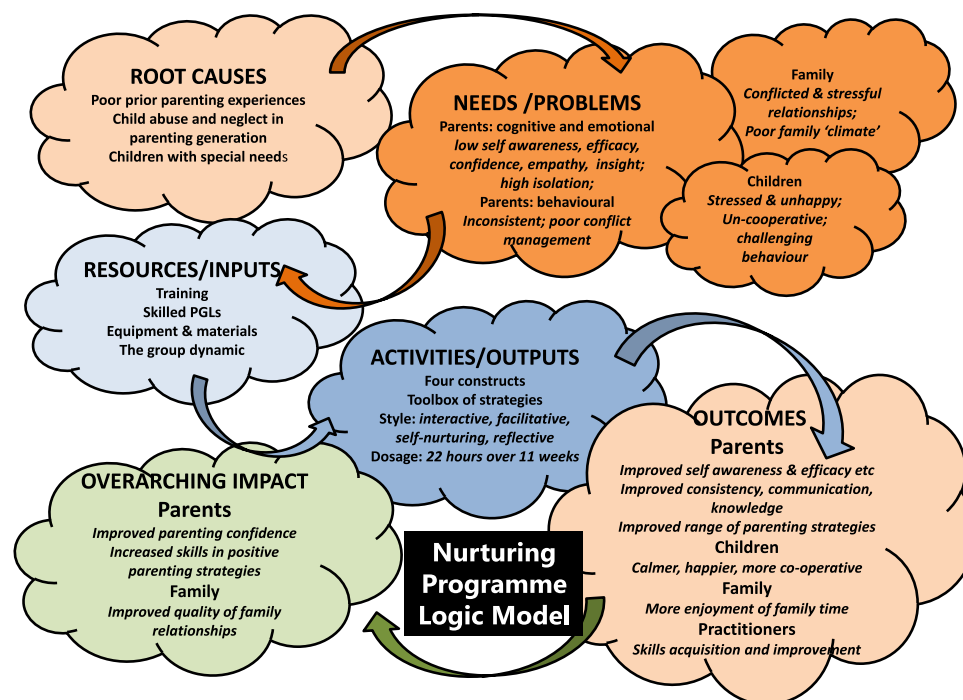


Fig. 4 Stage 1 Draft high level logic model for the FLNP-10. This Figure is covered by the Creative Commons Attribution 4.0 International License. Reproduced with permission of The Colebrooke Centre for Evidence and Implementation and Family Links

support, and so would be realistic given the particular type of intervention and implementation model employed by FLNP-10. Here, we were seeking broad endorsement (or rejection) of the key pillars and components of the programme, rather than definitive evidence of effectiveness/ineffectiveness, since the ultimate aim of the project was to develop a theory of change that would itself be thoroughly tested in due course.

The validation process consisted of a focused, purposive review by a researcher familiar with both the parenting support field and with review strategies, using a method specifically designed to map evidence directly to the questions raised by the theory of change, for example: How plausible is this theory? What evidence is there that *these* specific needs can be met through parenting support interventions delivered within *these* specific parameters? To what extent have other similar interventions of this type (group-based, community set, medium duration etc) achieved *these* kinds of outcomes? What alignment is there between the design and implementation method of FLNP-10 and other successful (or unsuccessful) initiatives?

To answer these questions, we worked through the Stage One theory model shown above in Fig. 3. We searched the international literature on parenting support ('parenting programmes/parent training and education/parenting support') by means of a range of scholarly databases and general purpose search engines, reference trails and hand searches, using key words and search terms derived from the model. Databases were used to identify key scholarly articles on parenting support interventions. Unpublished dissertations and non-peer reviewed material were excluded. We also identified a number of important text books dealing with theory and evidence in child development, parenting support and approaches to supporting parents through intervention.

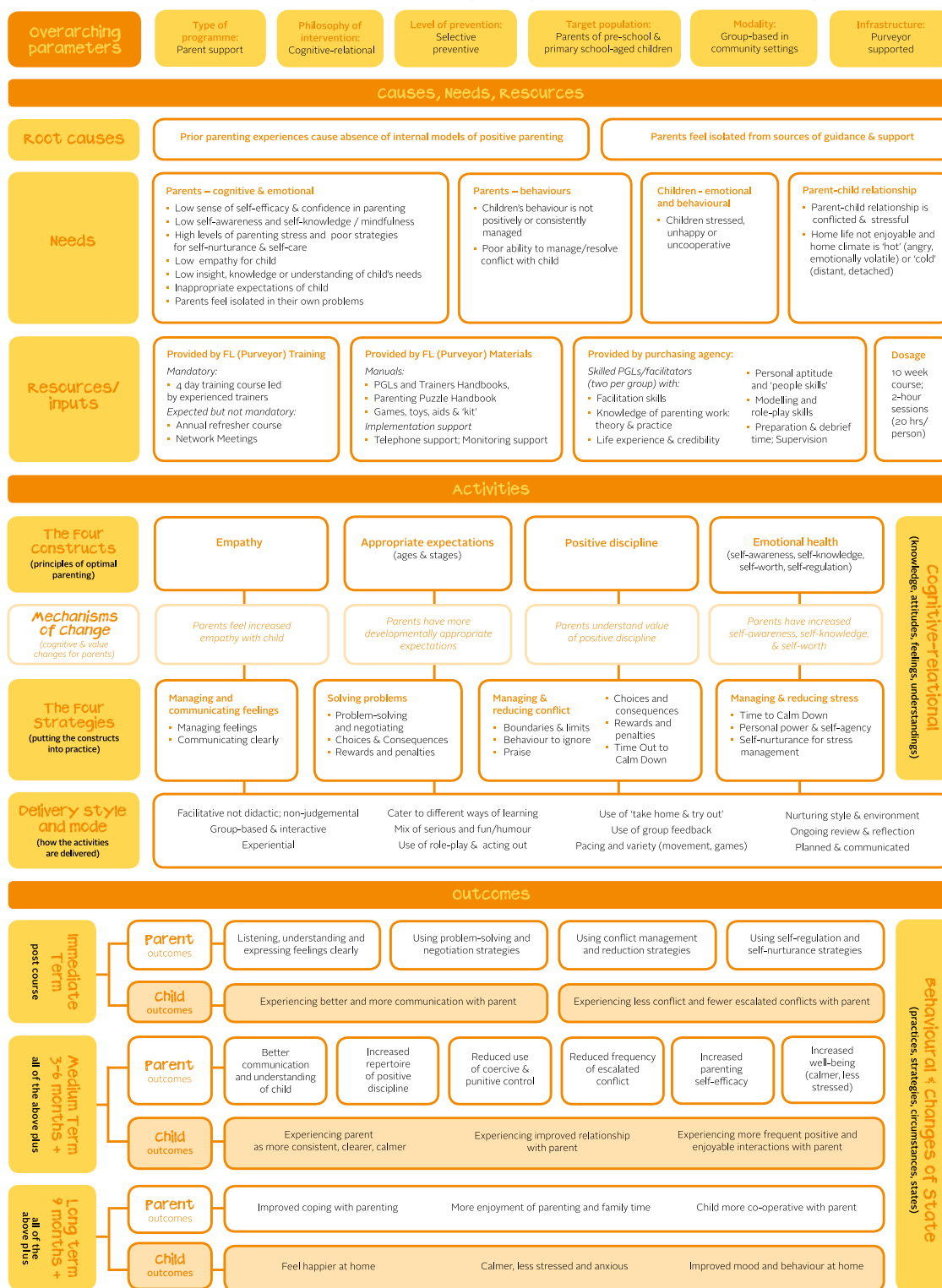
The key terms were selected to reflect the constructs, strategies, activities, delivery modality of the FLNP-10 and its potential outcomes as outlined at that stage in the Fig. 3 theory of change model. This process was dynamic and to some extent emergent. For example, key parameters (at that stage of the project) had not

yet been clearly identified, and as the search process unfolded, these and other lines of enquiry emerged as particularly important to follow. This method rather than more formal 'systematic' review method was used to ensure that the validation process remained cost-effective and proportionate to the project as a whole; it also allowed us to follow up specific issues of significance to FLNP implementation rather than to attempt to cover the entire, voluminous literature on effective parenting support, much of which would not have been relevant to our purpose (and see Greenhalgh et al., 2018, for a critique of systematic reviews).

We then reworked the model in consultation with new and original members of the content expert group to arrive at the final model, which is shown in Fig. 5.

Results and insights from the validation process. The validation process prompted fresh thinking about the programme theory when coupled to the emerging change model. Through discussion, some of which was demanding, it resulted in a sharper articulation both of the implementation model and the core components and plausible outcomes from the programme. For example, certain aspects of the historic design of the programme were able to be scrutinised in the light of contemporary evidence and theory, captured in a narrative paper written by the implementation consultant for the programme purveyors (Ghate, 2016). Through discussion of this paper, a co-produced 're-working' of the Stage One theory of change emerged, as shown in Fig. 5. The purveyors eventually made two relatively substantial changes to historic programme design, one to targeted outcomes (child abuse prevention) as discussed above, and one to content (promotion of self-esteem, see below), as well as many important clarifications including a sharpened articulation of the philosophy of intervention, a more precise definition of the target population, a firmer articulation of the level of prevention and a clearer specification of the primary as opposed to secondary outcomes to be expected.

Family Links 10-week Nurturing Programme: Theory of change



Created by Family Links and the Colebrooke Centre, 2016.

Fig. 5 Stage 2 Final full synthesis theory of change model for the FLNP-10. This Figure is covered by the Creative Commons Attribution 4.0 International License. Reproduced with permission of The Colebrooke Centre for Evidence and Implementation and Family Links; © Colebrooke Centre for Evidence and Implementation and Family Links, all rights reserved

The ‘revisiting’ of the construct of self-esteem provides an example of how this work produced change in practice. Historically, the programme design had laid emphasis on parents’ self-esteem and its relationship to optimal parenting, and it was one of the original constructs on which the programme rested. But in the intervening years, new research evidence had emerged showing self-esteem to have a more complex and mixed relationship to outcomes from intervention than previously thought, and this was illuminated through the validation process (see Ghate, 2016). Consideration of this development with the content experts led to a broader re-framing of this construct for the contemporary FLNP-10 under the general term of ‘emotional health’, focusing less on valuing oneself and more on how parents’ self-beliefs and self-management might impact on their children’s wellbeing and development⁷.

Conclusions

We suggest that there is hope of levelling the playing field across a diversity of community-based social programmes. In human services policy it seems we are often willing to spend far more resource on innovation than on supporting and improving what we already have, yet wholesale innovation may not always be the only option, and sometimes, it can be downright disappointing. The work detailed in this paper shows how application of implementation science-informed techniques can help home-grown programmes to compete scientifically with their more glamorous cousins, the so-called EBPs, by strengthening their design, delivery, and general evaluability. It offers the promise of better use of existing resources. The example used for illustration has shown a generalisable method for how such a process may be undertaken, co-produced by implementation specialists and programme purveyors and drawing on the best available evidence. It utilised improvement techniques in the real world in a cost-effective and practical way.

Reflecting time and limited resources as well as ongoing learning, the process we undertook was not perfect: for example, we were not able to include detailed thinking about implementation outcomes although we knew them to be important, and our Stage Two visualisation (Fig. 5) still retains a linearity that does not fully reflect the reality of the on-the-ground, messy and emergent process of implementation. The systemic aspect of how the programme functions was also not well-integrated, although we also know this to be significant (Ghate, 2015). Its inclusion would have challenged our thinking even more and would have further opened up the theoretical model. These remain aspirations for future work.

The work was not easy. Open-mindedness and a fair degree of courage on the part of programme purveyors or commissioners is required to entertain what may feel like substantial changes to a well-established and familiar way of doing things. In this project, the purveyors made considerable changes, hopefully improvements, to the established programme as a result of the work. While this is inevitably a nerve-testing process, confidence is strengthened by having gone through two linked, robust and careful processes (articulation of the theory of change accompanied by validation against existing evidence), so that there is a clear ‘trail’ showing why and how decisions were taken and how they may be tested in future. It has been highly effective in helping the programme purveyors to scrutinise their work, feel more confident about it, and prepare for a future with hopefully improved evidence of effectiveness. The model provided by the theory of change should, for example, create a clearer blueprint for future evaluators, so that future outcome measurement more faithfully reflects the actual implementation model of the

programme rather than the intuitions and assumptions of external evaluators.

We gained some important insights about the timing of theory of change development during the project. A general insight from this work is that developing a theory of change—labour intensive though it can be—is not a one-time only activity: even those interventions that have one should be prepared to review and update it regularly. Theories of change should clearly be seen as dynamic rather than static, evolving over time as learning from implementation in practice accumulates. We also learned something more. In an ideal world, as the orthodox view has it, a theory of change would generally be developed a priori, guiding intervention design as it is underway and before a programme’s implementation model is fully determined, and then revisited at various junctures over time. But, perhaps counterintuitively, this project shows there is also value in developing (or re-developing) a theory of change carefully against the actual, ‘in practice’ implementation model post hoc, once delivery has commenced. It is never too late to start. Some element of validation or ‘evidence checking’ (Gough and Boaz, 2015) to ensure that the assumptions set out in a theory of change have in fact been validated in other similar programmes is also desirable and feasible, and may lead to all kinds of unanticipated developments in thinking. In a new programme, using such a process would help ensure that ineffective methods were not being replicated unwittingly; and an established programme, it can help to illuminate aspects of historic design that might have become ‘activity traps’ or that might no longer be supported by best evidence. It may also illuminate ways in which a programme has evolved, perhaps in undocumented ways, over time.

Overall, the practice of theory development and visualisation needs to evolve. Real world theories of change benefit from being visualised as dynamic and interactive rather than linear, with multiple pathways to end results, and multiple feedback loops and multi-directional flows to accommodate the real world complexity of most human behaviours, relationships and situations. They probably need to accommodate or allow for unknowns (or ‘emergence’) much more explicitly than they currently do. We found no examples in the existing literature of how one might construct such a model that still conforms to the certainties often required by programme funders and commissioners of social programmes and yet is not so complex that it becomes unfathomable. This may be a fruitful area for future collaborative development work for implementation and improvement scientists with practitioners and purveyors—especially if we want to find cost-effective ways to strengthen programmes and services that already exist as well as investing in innovative approaches.

Received: 9 January 2018 Accepted: 11 June 2018

Published online: 17 July 2018

Notes

¹ <https://familylinks.org.uk/the-nurturing-programme>

² Purveyors are individuals or organisations that operate in the role of external experts to a provider implementing a particular programme. They support organisations, systems and practitioners in striving to adopt and implement that programme with fidelity and effectiveness (Oosthuizen and Louw, 2013; Franks and Bory, 2015)

³ <http://www.nurturingparenting.com/>

⁴ Campbell’s Law (Campbell, 1979) famously asserts that quantitative measurement—with all the consequences that are typically attached to it—is apt to distort and corrupt the very social processes that are being measured. In other words, the measurement rather than the delivery of effective services becomes the focus.

⁵ Some authors (e.g., Proctor et al, 2011) have used the term ‘implementation outcomes’ to describe as attributes, features or outputs of the implementation process, including

acceptability to users, cost and feasibility. This is not the sense in which we use the term here.

6 See for example <http://www.theoryofchange.org/library/toc-examples/>

7 See: <https://www.emotionalhealthatwork.org.uk/>

References

- All Party Parliamentary Group for Children (2017) No Good Options: Report of the Inquiry into children's social care in England, March 2017. NCB, London, <https://www.ncb.org.uk/sites/default/files/uploads/No%20Good%20Options%20Report%20final.pdf>
- Baker EL (2010) Taking programs to scale: a phased approach to expanding proven interventions. *Public Health Manag Pract* 16(3):264–269
- Baker M (2016) Is there a crisis of reproducibility? *Nature* 533:452–452
- Bavolek S (2000) The Nurturing Parenting Programs OJJDP Juvenile Justice Bulletin Nov. 2000 US Dept. of Justice: OJJDP <https://www.ncjrs.gov/pdffiles1/ojjdp/172848.pdf>
- Biehler N et al. (2010) The Care Placements Evaluation (CaPE) Evaluation of Multidimensional Treatment Foster Care for Adolescents (MTFC-A). London, Department for Education DfE RR 194
- Blase K, Fixsen D (2013) Core intervention components: identifying and operationalizing what makes programs work. US Dept of Health and Human Services, Washington DC, http://aspe.hhs.gov/hsp/13/KeyIssuesforChildrenYouth/CoreIntervention/rb_CoreIntervention.cfm ASPE Research Brief February 2013[on-line]
- Bradach J (2003) Going to Scale: The Challenge of Replicating Social Programs. *Stanford Social Innovation Review*, Spring pp19–25, https://ssir.org/articles/entry/going_to_scale
- Braithwaite J et al. (2018) When complexity science meets implementation science: a theoretical and empirical analysis of system change. *BMC Medicine* 16 (63) <https://doi.org/10.1186/s12916-018-1057-z>
- Campbell DT (1979) Assessing the impact of planned social change Evaluation and Program Planning 2(1):67–90. [https://doi.org/10.1016/0149-7189\(79\)90048-X](https://doi.org/10.1016/0149-7189(79)90048-X)
- Cartwright N (2011) The art of medicine: a philosopher's view of the long road from RCTs to effectiveness. *Lancet* 377 April 2011 1400–1401
- Cartwright N, Hardie J (2012) Evidence-based policy: A practical guide to doing it better. Oxford University Press, Oxford
- Chambers DA, Glasgow RE, Stange KC (2013) The dynamic sustainability framework: addressing the paradox of sustainability amid ongoing change. *Implement Sci* 8(117) <https://doi.org/10.1186/1748-5908-8-117>
- Cottrell DJ et al. (2018) Effectiveness of systemic family therapy versus treatment as usual for young people after self-harm: a pragmatic, phase 3, multicentre, randomised controlled trial. *Lancet Psychiatry* 5(3):201–2016
- Crossley S (2015) The troubled families programme: the perfect social policy? Centre for Crime and Justice Briefing, 13 Nov 2015 <https://www.crimeandjustice.org.uk/publications/troubled-families-programme-perfect-social-policy>
- Davis R et al. (2015) Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health Psychol Rev* 9(3):322–344
- Davies P, Walker A, Grimshaw JM (2010) A systematic review of the use of theory in guideline dissemination and implementation strategies and interpretation of the results in rigorous evaluations. *Implement Sci* 5(14) <https://doi.org/10.1186/1748-5908-5-4>
- Dick AJ, Rich W, Waters T (2016) Prison vocational education policy in the United States: A critical perspective on evidence-based reform. Palgrave Macmillan, New York
- Durose C et al. (2017) Generating 'good enough' evidence for co-production. *Evid Policy* 13(1):135–151. <https://doi.org/10.1332/174426415X14440619792955>
- Edmiston D, Nicholls A (2018) Social Impact Bonds: the role of private capital in outcome-based commissioning. *J Social Policy* 47(1):57–76. <https://doi.org/10.1017/S0047279417000125>
- Fixsen DL et al. (2005) Implementation research: A synthesis of the literature. University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network, Tampa, FL, <http://nirn.fpg.unc.edu/resources/implementation-research-synthesis-literature> (FMHI Publication #231)
- Fonagy P, et al. (2018) Multisystemic therapy versus management as usual in the treatment of adolescent antisocial behaviour (START): a pragmatic, randomised controlled, superiority trial. *The Lancet Psychiatry* 5(2):119–133
- Forrester D et al. (2018) A randomized controlled trial of training in Motivational Interviewing for child protection. *Children and Youth Services Review* 12 Feb 2018. <https://doi.org/10.1016/j.childyouth.2018.02.014> <https://www.sciencedirect.com/science/article/pii/S0190740917308915>
- Franks RP, Bory CT (2015) Who supports the successful implementation and sustainability of evidence-based practices? Defining and understanding the roles of intermediary and purveyor organizations. *New Dir Child Adolesc Dev* 149:41–56 <https://onlinelibrary.wiley.com/doi/abs/10.1002/cad.20112>
- Ghate D (2015) From programs to systems: deploying implementation science and practice for sustained real-world effectiveness in services for children and families. *J Clin Child Adolesc Psychol* 45(6): Good Enough? Interventions For Child Mental Health: From Adoption To Adaptation—From Programs To Systems 10.1080/15374416.2015.1077449 <http://www.tandfonline.com/doi/full/10.1080/15374416.2015.1077449>
- Ghate D (2016) The Family Links 10-Week Nurturing Programme: developing a theory of change for an evidence-supported design <https://familylinks.org.uk/why-it-works#10-week-Nurturing-Programme>
- Gough D, Boaz A (2015) Evidence checks and the dynamic interactive and time- and context-dependent nature of research (Editorial). *Evid Policy* 11(1):3–6
- Grant R, Hood R (2017) Complex systems, explanation and policy: implications of the crisis of replication for public health research. *Crit Public Health* pp 1–8 <https://doi.org/10.1080/09581596.2017.1282603>
- Greenhalgh T, Thorne S, Malterud K (2018) Time to change the spurious hierarchy of systematic over narrative reviews? *Eur J. Clinical Invest.* <https://doi.org/10.1111/eci.12931>
- Hammersley M (2005) Is the evidence-based practice movement doing more harm than good? Reflections on Iain Chalmers' case for research-based policy making and practice. *Evid Policy* 1(1):85–100
- Haynes A et al. (2016) Figuring out Fidelity: a worked example of the methods used to identify, critique and revise the essential elements of a contextualised intervention in health policy agencies. *Implement Sci* 11(23) <https://doi.org/10.1186/s13012-016-0378-6>
- Hawe P (2015) Lessons from complex interventions to improve health. *Annu Rev Public Health* 36:307–323
- Hawe P, Shiell A, Riley T (2004) Complex interventions: how "out of control" can a randomised controlled trial be? *Br Med J* 328(7455):1561–1563
- Huberman M (1999) The mind is its own place: the influence of sustained interactivity with practitioners on educational researchers. *Harv Educ Rev* 69 (3):289–319
- Humayun S et al. (2017) Randomised controlled trial of Functional Family Therapy for offending and antisocial behaviour in UK youth. *J Child Psychol Psychiatry* 58(9):1023–1032. <https://doi.org/10.1111/jcpp.12743>
- Kainz K, Metz A (2016) Causal thinking for embedded, integrated implementation research. *Evidence Policy* <https://doi.org/10.1332/174426416X14779418584665>
- Kaminsky JW et al. (2008) A meta-analytic review of components associated with parent training program effectiveness. *J Abnorm Child Psychol* 36:567–589
- Kaplan SA, Garrett KE (2005) The use of logic models by community-based initiatives. *Eval Program Plan* 28:167–172
- Little M et al. (2012) The impact of three evidence-based programmes delivered in public systems in Birmingham UK. *Int J Confl Violence* 6(2):260–272
- Melhuish E et al. (2008) Effects of fully-established Sure Start Local Programmes on 3-year-old children and their families living in England: a quasi-experimental observational study. *Lancet* 372(9650):1641–1647. [https://doi.org/10.1016/S0140-6736\(08\)61687-6](https://doi.org/10.1016/S0140-6736(08)61687-6)
- Moore G et al. (2015) Process evaluation of complex interventions: Medical Research Council guidance A report prepared on behalf of the MRC Population Health Science Network. London: Institute of Education <https://www.mrc.ac.uk/documents/pdf/mrc-phsrn-process-evaluation-guidance-final/>
- Moran P, Ghate D (2013) Development of a single overarching measure of impact for Home-Start: a feasibility study. The Colebrooke Centre, London, http://www.cevi.org.uk/docs/Impact_Measure_Report.pdf
- Morpeth L et al. (2017) The effectiveness of the Incredible Years pre-school parenting programme in the United Kingdom: a pragmatic randomised controlled trial. *Child Care Pract* 23(2):141–161. <https://doi.org/10.1080/13575279.2016.1264366>
- Mountford A, Darton S (2013) Family Links response to Simkiss et al. (2013) Letter to the BMJ. <http://bmjopen.bmj.com/content/3/8/e002851.responses#family-links-response-to-effectiveness-and-cost-effectiveness-of-a-universal-parenting-skills-programme-in-deprived-communities-multicentre-randomised-controlled-trial---simkiss-et-al>
- Mowles C (2014) Complex, but not quite complex enough: the turn to complexity sciences in evaluation scholarship. *Evaluation* 20(2):160–175
- Mulgan G (2016) What's wrong with theories of change? Alliance for Useful Evidence Blog, 6th Sept 2016 <http://www.nesta.org.uk/blog/whats-wrong-theories-change>
- Oosthuizen C, Louw J (2013) Developing program theory for purveyor programs Implementation. *Science* 8:23, <http://www.implementationscience.com/content/8/1/23>
- Pawson R (2013) The science of evaluation: A realist manifesto. Sage Publications, London
- Pawson R, Tilley N (1997) Realistic evaluation. Sage Publications, London
- Pearce W, Raman S, Turner A (2015) Randomised trials in context: practical problems and social aspects of evidence-based medicine and policy. *Trials* 16:394. <https://doi.org/10.1186/s13063-015-0917-5>

- Pfitzer M, Bocksette V, Stamp M (2013) Innovating for shared value. *Harvard Business Review* <http://hbr.org/2013/09/innovating-for-shared-value/ar/pr>
- Proctor E et al. (2011) Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Heath* 38:65–76. <https://doi.org/10.1007/s10488-010-0319-7>
- Renger R, Hurley C (2006) From theory to practice: lessons learned in the application of the ATM approach to developing logic models. *Eval Program Plan* 29:106–119
- Renger R, Titcomb A (2002) A three-step approach to teaching logic models. *Am J Eval* 23(4):493–503
- Rittel HWJ, Webber MM (1973) Dilemmas in a general theory of planning. *Policy Sci* 4(2):155–169
- Robling M et al. (2015) Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (Building Blocks): a pragmatic randomised controlled trial. *Lancet*. [https://doi.org/10.1016/S0140-6736\(15\)00392-X](https://doi.org/10.1016/S0140-6736(15)00392-X)
- Rogers PJ (2008) Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation* 14(1):29–48. <https://doi.org/10.1177/1356389007084674>
- Schneider F et al. (2017) Impacts of social learning in transformative research Blog for Integration and Implementation Insights, 16th May 2017 <https://i2insights.org/2017/05/16/social-learning-impacts/>
- Simkiss DE et al. (2013) Effectiveness and cost-effectiveness of a universal parenting skills programme in deprived communities: a multicentre randomised controlled trial. *BMJ Open* 2013(3):e002851. <https://doi.org/10.1136/bmjopen-2013-002851>
- Stewart-Brown S et al. (2011) Should randomised controlled trials be the “gold standard” for research on preventive interventions for children? *J Children’s Serv* 6(4):288–235
- UK Implementation Society (2017) <https://www.ukimplementation.org.uk/>
- Utting D, Monteiro H, Ghate D (2007) Interventions for children at risk of developing anti-social personality disorder. Policy Research Bureau and Cabinet office, London, <http://www.prb.org.uk/publications/P182%20and%20P188%20Report.pdf>
- Weiss CH (1995) Nothing as practical as good theory: exploring theory-based evaluation for comprehensive community initiatives for children and families. In: Connell JP, Kubisch AC, Schorr LB and Weiss CH (eds) *New approaches to evaluating community initiatives*. Aspen Inst, Washington, pp. 65–69
- Weiss CH (1997) Theory-based evaluation: past, present and future. *New Dir Eval* 76:41–55. Winter 1997
- Whetten DA (1989) What constitutes a theoretical contribution? *Acad Manag Rev* 14(4):490–495. <http://www.jstor.org/stable/258554>
- Wong G et al. (2012) Realist methods in medical education research: what are they and what can they contribute? *Med Educ* 46(1):89–96 <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2923.2011.04045.x/full>

Data availability

Data sharing is not applicable to this paper as no datasets were analysed or generated.

Acknowledgements

The research on which this article is based was funded by Family Links. I thank Professor John Coleman for helpful comments on earlier elements of the work.

Additional information

Competing interests: The author declares no competing interests.

Reprints and permission information is available online at <http://www.nature.com/reprints>

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018