The third sector delivery of social services:

Needs, effectiveness and strategies for improvement



Anders Malthe Bach-Mortensen

Student number 848999

Wolfson College

Supervisors: Paul Montgomery & Jane Barlow

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Social Intervention

Table of Contents

Thesis Abstract 5

Chapter 1: Introduction 9

Research questions and objectives 11

Thesis outline 11

Chapter 2: Background 13

Defining the third sector and its component organisations 13

The definition applied in this thesis 15

Mixed provision of social services – blurring of sectors? 16

'Distinctiveness' of the third sector 19

Towards a (clearer) typology of sector 21

Unique characteristics influencing different types of service providers 23

A revised framework of comparative advantage and disadvantage 27

Evidence-based practice and policy-making 30

Evidence based practice 30

Evidence-based policy-making 32

Evidence-based policy making in the UK 33

How this applies to third sector service provisions 34

Evaluation and implementation among third sector providers of public services 35

Policy assumptions underlying the opening of public services 39

Papers 1 & 2: What issues do TSOs experience when evaluating and implementing their services? 39

Paper 3: Quality variation among for-profit, public, and third sector social care providers 41

Conclusion 44

Bibliography 44

Chapter 3: What are the barriers and facilitators for third sector organisations (non-profits) to evaluate their services? A systematic review 51

Abstract 52

Background 54

Objectives 57

Methods 57

Search strategy 57

Study selection 58

Screening and data-extraction 59

Data synthesis 59

Quality assessment 60

Results 60

Characteristics of included studies 61

Samples of TSOs in included studies 62

Size of TSOs 62

Purpose of TSOs 63

Type of evaluation 63

Quality of included studies 64

Identified barriers and facilitators 64

Weight of factors 65

Themes 65

Robustness of themes 70

Sensitivity analysis 70

Discussion of findings 71

Summary of findings 71

Implications of findings 71
Ways forward 73
Limitations of this study 75
Future research 76
Conclusions 77
References 78
Chapter 4: Barriers and facilitators to implementing evidence-based interventions among third sector organisations: a systematic review 85
Abstract 86
Background 88
What has been done? 90
Objectives 91
Methods 92
Search strategy 92
Selection criteria 92
Data extraction 93
Data synthesis 94
Quality appraisal 95
Results 95
Characteristics of included studies 97
Sample size of included studies 97
Types of organisations 98
Types of EBIs 98
Quality of studies 100
Barriers and facilitators 100
Thematic analysis 101
Robustness of findings 105
Recommendations for policy and practice 109
Discussion 110
Summary of findings 110
Implications for policy and practice 111
Ways forward 113
Limitations 115
Further research 115
Conclusion 116
Bibliography 117
Chapter 5: Does sector matter for the quality of care services? A secondary analysis of social care services regulated by the Care Inspectorate in Scotland 126
Abstract 127
Introduction 129
Objectives 131
Data: The Care Inspectorate 131
Methods 133
Explanatory variables 134
Dependent variables 135
Results 135
Descriptive statistics 135
Predicted probabilities 138
Difference between for-profit, public and third sector organisations 139
Differences between public and third sector providers 1

Risk of services 1
Complaints and requirements 2
Other explanatory factors 3
Sub-analyses 3
Discussion 4
Strengths and limitations 5
Implications for policy and practice 6
Conclusion 7
Bibliography 8
Chapter 6: Discussion and recommendations 11
Summary of findings 11
Paper 1: Barriers and facilitators to evaluation among third sector organisations 11
Paper 2: Barriers and facilitators to implementing evidence based interventions among third sector organisations 12
Paper 3: Quality variation among for-profit, public, and third sector social care providers 13
Limitations 14
Paper specific limitations 14
General limitations of thesis 16
Recommendations 17
Recommendation 1: Incentivise an evaluation culture of learning 17
Recommendation 2: Design clearly defined interventions that can be adapted to local contexts and populations 22
Recommendation 3: Policy and regulation should facilitate incentives that focus on quality rather than costs 25
Future research directions 27
Further research on third sector evaluation 27
Further research on quasi-market provisions 28
Conclusions 29
Bibliography 30
Appendices 34
Appendix A (Chapter 3) 35
Appendix A1: Adapted checklist for quality appraisal of cross-sectional findings 35
Appendix A2: Adapted CASP and JBI checklist 37
Appendix A3: Top 3 cited barriers excluding the six studies of low quality. 38
Appendix A4: Themes excluding low quality studies. 39
Appendix A5: Search strategy for included databases 41
Appendix A6: PRISMA diagram 45
Appendix B (Chapter 4) 48
Appendix B1: Prisma Diagram 48
Appendix B2: Search strategy 50
Appendix B3: Extracted quotes on barriers 51
Appendix B4: Extracted quotes on facilitators 75
Appendix B5: Adapted checklist for quality appraisal of cross-sectional findings 91
Appendix B6: Adapted CASP and JBI checklist 93
Appendix B7: Less significant recommendation themes 94
Appendix B8: Quality Appraisals of qualitative studies 97
Appendix B9: Quality Appraisals of Quantitative Studies 99
Appendix C (Chapter 5) 101
Appendix C1: descriptive information and predicted probabilities 101
Appendix C2: Regression coefficients of all models 107
Appendix C3: Predicted probabilities across individual service types 119

Thesis Abstract

Background

The third sector is expanding and becoming a bigger provider of public, social and health care services. However, little research has been undertaken to date to evidence the *ability* of third sector organisations (TSOs) as substitute health and social care providers to work effectively and in accordance with evidence-based practice (EBP). This entails having established processes for monitoring service delivery and outcomes, and including relevant stakeholders with the best available evidence in decision-making processes. Whilst the principles of EBP are gradually becoming an accepted part of practice in health and social care settings, little is known about how these principles have transferred into the third sector.

To meet the growing demand on health and social care as a result of changes such as ageing populations, such services are increasingly delivered in quasi-markets, which are markets where public, for-profit, and third sector providers compete for service contracts financed by the state. Research convention and theories on the typology of sector have long contended that organisations across sector are driven by different incentives, but it remains unclear how, if at all, sector specific traits influence service *quality*. Further, little research has investigated the implications associated with the use of quasi-markets and whether certain types of service providers deliver systematically higher value than others.

Objectives

This thesis aims to inform these areas by conducting two systematic reviews, which will be operationalised as an evaluation of the factors influencing the ability of TSOs to (1) evaluate their activities and (2) implement evidence-based interventions. (3) The thesis then tests whether quality-related outcomes vary across for-profit, public and third sector organisations in the context of social care organisations in Scotland. These three pieces of research have been

developed and written as stand-alone journal articles for publication, bookended by chapters that will engage in a broad discussion of the core issues on the topic and of the core research questions. Together, these three papers provide an important and novel contribution to the existing literature by identifying, appraising, and synthesising existing research on the evaluation and delivery issues of TSOs, while also empirically exploring whether quality outcomes differ across public, for-profit, and third sector social care providers.

Methods

A variety of methods are employed throughout the thesis. Papers 1 and 2 are systematic reviews conducted following best practice and in line with PRISMA guidelines. Both reviews included systematically searching key databases using relevant search strings, contacting experts, and reviewing websites. All included literature was quality appraised and synthesised using thematic analysis. Paper 3 employed generalised ordered logit models utilising data from the Care Inspectorate (the independent regulator of care services in Scotland) to test whether service-related outcomes vary across for-profit, public and third sector organisations. All papers were reported following best practice guidelines.

Results

Paper 1 identified the main barriers for TSOs undertaking evaluation to be related to the lack of financial resources, lack of technical capability and evaluation literacy, and challenges around identifying relevant evaluation systems and outcome indicators. Key facilitating factors involved securing the appropriate support and having an organisational culture that supports evaluation, in addition to the motivation to be accountable to stakeholders. This paper has been published in Systematic Reviews [1]. Paper 2, which focused on factors influencing the ability of TSOs to implement evidence-based interventions (EBIs), identified resource limitations, in particular staff and finance, to be the most reported barrier to TSOs implementing EBIs. Organisational *culture*,

including factors such as alignment between the mission of the TSO and EBI, and support/prioritisation of the implementation process, were the most reported facilitators. This paper has been published in Implementation Science [2]. Paper 3 found that public and third sector social care providers were generally more likely to be rated of high quality and of lower risk, compared to for-profit organisations, which were more likely to be rated of adequate quality and of higher risk. This paper has been published in BMJ Open [3].

Discussion

The findings of these three studies suggest that from a regulation perspective, the large range of different activities within the third sector makes it challenging to develop policies that are appropriate to all types of TSOs. While policy-makers and academics continue to emphasise the importance of the third sector in alleviating social issues, TSOs are subject to little and insufficient guidance and support on how to evaluate and implement their services in accordance with best practice. This lack of clarity and support regarding how TSOs should adhere to best practice is problematic and may largely be at the expense of the beneficiaries; most of whom include vulnerable population groups.

The findings presented in this thesis constitute a novel contribution to the knowledge-base regarding the current state of affairs in the third sector delivery of health and social care services, and also informs theories on the typology of sector. However, the findings should be interpreted with caution as a result of the study limitations, which are discussed in full. Further, the practical and policy implications identified in each paper provide an important resource for key stakeholders in terms of future strategies for regulation and improved guidance/support. As such, the findings do not just have academic relevance in terms of improving the evidence-base with regard to existing evaluation, implementation and quality practices among TSOs, but also

inform practice in terms of current debates about how the needs of TSOs can be met going forward.

References

1. Bach-Mortensen AM, Montgomery P. What are the barriers and facilitators for third sector organisations (non-profits) to evaluate their services? A systematic review. Syst Rev. 2018;7:13. doi:10.1186/s13643-018-0681-1.

2. Bach-Mortensen AM, Lange BCL, Montgomery P. Barriers and facilitators to implementing evidence-based interventions among third sector organisations: a systematic review. Implement Sci. 2018;13:103. doi:10.1186/s13012-018-0789-7.

3. Bach-Mortensen AM, Montgomery P. Does sector matter for the quality of care services? A secondary analysis of social care services regulated by the Care Inspectorate in Scotland. BMJ Open. 2019;9:e022975. doi:10.1136/bmjopen-2018-022975.

Chapter 1: Introduction

In most welfare states, social services, such as elderly and child care, are delivered by both for-profit, public, and third sector providers. These services are increasingly commissioned under 'quasi-markets', in which providers from all sectors compete for public service contracts [1, 2]. Currently, most adult residential services and children homes in the UK are for-profit (60% to 80%), followed by public (10% to 20%) and third sector providers (15% to 20%)[1]. This can be thought of as a positive development following the rationale that such market structure enables service contracts to be allocated to the most capable provider regardless of sector. However, despite the intuitive rationale that open bidding facilitates better services and more consumer choice, the marketization of social care has been criticized for relying on principles of new public management, which prioritises cost-effectiveness above service quality [3–5].

The third sector, in particular, is a significant and growing provider of public, health and social care services. However, while policy-makers continue to emphasise the increased inclusion of the third sector in the delivery of such services [6–9], there is limited evidence of the *effectiveness* of the third sector substitution of public and social services [10]. Indeed, there is little research addressing the *ability* of the third sector to deliver public services in accordance with 'best practice', thus making it unclear whether third sector services are safe and effective.

Evidence-based practice (EBP) is becoming increasingly integrated in today's political discourse [6, 10–12], and multiple governments have committed themselves to so-called 'evidence-based policy-making'. This discourse is, for example, illustrated in the recent White Paper 'Open Public Services', which emphasises the use of evidence and evaluation in the delivery of public services

[1] These proportions vary according to provider types. Nursing homes: CQC and Care Inspectorate data. Children homes: Ofsted.

[6]. Further examples include the creation of the UK 'What Works' centres, and the initiation of the 'Evidence-Based Policy Making Commission Act of 2016' in the US [13]. However, despite the proclaimed political *willingness* to provide services based on the best available evidence, the main legal requirement for UK TSOs is that they report on their governance and expenditure [14].

The Charity Commission in England and Wales requires charities to report on their 'achievements', but does not specify what constitutes an 'achievement' and similarly fails to define how achievements from different activities should be evaluated. Another example of the apparent willingness to enforce evaluation without having a workable conceptualisation of how to measure and report on the impact of public services is exemplified by the Social Value Act of 2012, which requires public authorities to "consider" the social impact of the services they provide and commission. However, a recent report evaluating the effect of the Social Value Act found that the lack of consensus and consistency regarding the notion of 'social impact' obstructed public authorities from providing consistent measures on the impact of the services they provide and commission [15].

It might be argued that evaluation takes resources away from the actual activities of TSOs, effectively 'wasting' valuable resources that could be spent on charitable services. For example, research illustrates that evaluation by TSOs is often driven by the need to meet funder requirements, as opposed to improving the performance of services [16, 17]. However, the case for enforcing service evaluation rests on important arguments, especially when working with social interventions and vulnerable population groups. Most centrally, evaluation and monitoring should be embraced as an endeavour to ensure *service effectiveness* and to fulfil the Hippocratic Oath, *first do no harm* [18, 19]. Considering that most interventions have the potential to do harm to service-users [20–23], it is imperative that regulators and commissioners consider those risks

as part of their criteria to service providers. Further, not having sound evaluation practices in place also involves lost opportunity costs, as evaluation data can be utilised as a management tool that refines and improves services. It is thus imperative that future efforts are made to establish a common framework and criteria regarding concepts such as 'social impact' that is enforceable by regulators and commissioners and that follows best practice.

Research questions and objectives

Third sector activities are increasing, but there is limited research addressing the ability of TSOs to deliver public and social services in accordance with best practice. This thesis seeks to improve the evidence regarding the ability of the third sector to provide evidence-based services and to investigate whether the quality of services varies across different providers (third, for-profit and public sector). In so doing, it will ask the following research questions:

(1) What factors influence the ability of TSOs to evaluate their services (paper 1)?
(2) What factors influence the ability of TSOs to implement evidence-based interventions (paper 2)?
(3) Does the quality of social care services vary according to ownership (paper 3)?

These questions will be addressed via a 'stand-alone papers' format thesis and will thus be based on peer-reviewed academic publications.

Thesis outline

First, this thesis defines the main concepts explored in the research, conceptualising what characteristics influence for-profit, public, and third sector social care providers, and elaborating on the political discourse surrounding the mixed delivery of public services. Second, the results

from a systematic review on the evaluation issues of TSOs, operationalised as the barriers and facilitators that TSOs experience in evaluating their activities, are reported. Third, the thesis reports on findings from an additional systematic review, which focuses on the barriers and facilitators that TSOs experience when implementing evidence-based interventions. Fourth, the thesis tests whether service quality varies across for-profit, public and third sector social care providers in Scotland. Last, the thesis provides a discussion of the overarching recommendations and limitations identified in the empirical chapters.

Chapter 2: Background

Defining the third sector and its component organisations

The 'third sector' is a vague and inconsistently applied term, which encompasses a range of different types of organisation, such as 'voluntary organisations', 'community-based organisations', 'non-profits' and 'charities', etc. [24–26]. Having one definition to cover all organisations within the third sector may be unsatisfactory, considering the vast and diverse range of activities that fall within this sector. That is, third sector organisations (TSOs) can undertake diverse types of provision ranging from religious clubs, housing, HIV, children's sport clubs, hospices, aid and development, mental health and many other types of activities. To complicate things further, there is usually a divide between the conceptualisation of TSOs in Europe and in the US [25].

It is therefore hardly surprising that scholars continue to struggle to find a universally accepted definition of the third sector [24, 27]. However, most academics tend to draw on the definition by Anheier and Salamon made under the John Hopkins Center for Civil Societies (1997). This definition emphasises five characterising traits of a third sector organisation, i.e. that they must be: (1) a formal organisation, (2) privately owned and independent from the government, (3) non-profit distributing, (4) self-governing, and (5) benefitting from voluntary activities. This conceptualisation is referred to as the *structural-operational* definition, and its applicability has been tested in a range of comparative cases [28, 29]. The main purpose of this definition is to provide a universal and workable definition of organisations falling within the third sector.

Whilst the universality of this definition has been subject to some criticisms (see e.g. Morris 2000 [30]), it continues to be one of the main applied definitions in practice [31, 32]. Importantly, one

might argue that the heterogeneity of third sector organisations makes the ambition to develop a universal definition a Sisyphean pursuit, which will never satisfy the diverse set of characteristics of the organisations that are in the sector. To illustrate the realm of the third sector, the below model may serve as a helpful conceptualisation of how the third sector relates to the for-profit and public sectors.

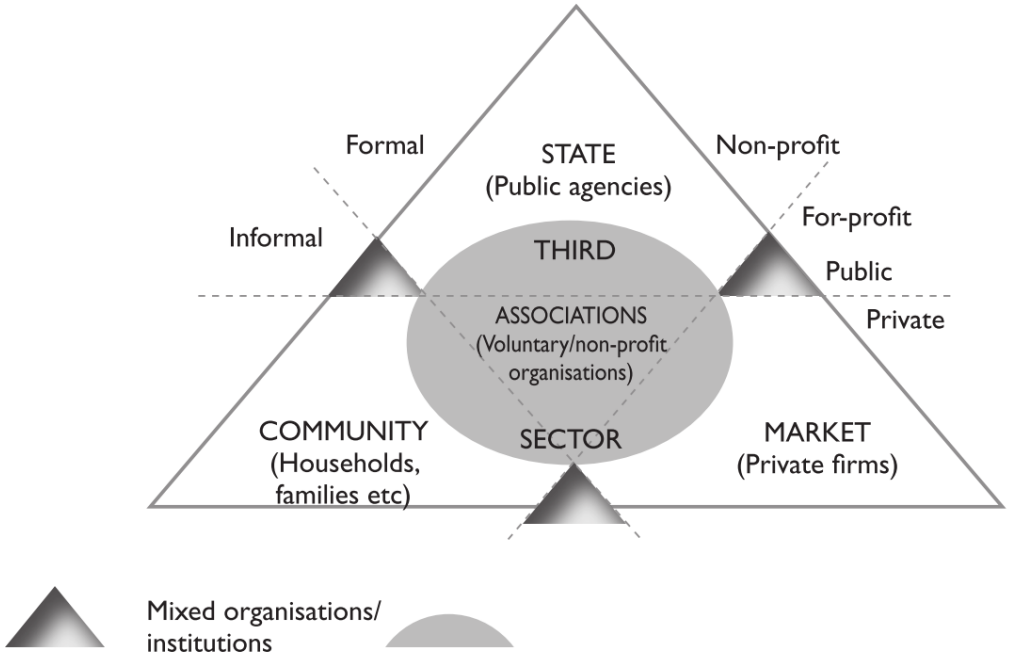


Figure 1: The welfare mix (adapted from Evers and Laville 2006)

This model essentially suggests that one can split up the state, the market and the community sector, and that the third sector includes elements from all these sectors. Organisations within the third sector thus include characteristics from some or all those sectors, often making them hybrid organisations [33]. In fact, it is currently debated whether the demarcation of the sectors provided by, for example, the 'welfare mix model' (illustrated above) is sufficient to account for the diverse types of TSOs [34], considering that the third sector entails a lot of hybrid organisations that contain characteristics from either or both the for-profit and public sector.

However, recent attempts to define the third sector suggested caution when attempting to identify a universal definition of the third sector, as this might operate as a barrier to conducting research on the sector [24, 35]. Rather, the Third Sector Research Centre[2] at Birmingham University advocates for a positive and inclusive definition. Specifically, the centre argues that all organisations that are active within the third sector can be considered TSOs:

> "We conceptualise the third sector as a space of organisational activity located between the state, market and private familial spheres comprising a diversity of organisational types including charities, social enterprises, faith, community and grassroots groups." (Rees and Mullins 2016: 3)

One might criticise this conceptualisation for being too broad in its reach, effectively failing to account for important subsector differences of TSOs. However, considering the heterogeneity of organisations constituting the sector, one might argue that splitting the sector into too many parts will make the overall sector concept less useful in researching the organisations within it [34]. Moreover, Alcock (2010) argues that the current political discourse of the 'third sector' is a product of what he refers to as 'strategic unity', in which heterogeneous organisations such as social enterprises and different types of TSOs are coming together under the same banner. Doing so offers a "marriage of convenience" [24], which can be thought to strengthen the overall sector, despite the intrinsic differences of the included organisations.

The definition applied in this thesis

This thesis will apply the, by now, more commonly applied term *third sector organisation* (TSO), which is becoming a more universally used denotation of organisations within the third sector

[2] http://www.birmingham.ac.uk/generic/tsrc/index.aspx

[35]. In so doing, the thesis will draw on the structural-operational definition by Anheier and Salamon (1997) as a working definition. However, research conducted on TSOs rarely provides sufficient information on whether the included TSOs fulfil the five conditions of that definition. Therefore, this thesis will also allow for a more inclusive definition similar to the one provided by the Third Sector Research Centre.

Thus, this thesis will embrace a generic and inclusive definition of the third sector, in which any organisation working within it, or claiming to be working within it, will be considered as a TSO. Doing so allows for addressing the sector more broadly and embracing the increasing heterogeneity of the third sector. However, an inclusive definition comes at the cost of undermining the specificity in terms of the subsectors. To address this shortcoming, the thesis will endeavour to describe and control for organisational heterogeneity (when possible) in its analysis of TSOs.

Mixed provision of social services – blurring of sectors?

Due to the privatisation of state service provisions, the third sector is becoming a growing provider of public and social services [10, 29, 36]. The 'opening' of public services in the UK, but also in other OECD countries has increased the market for the third sector and thus facilitated growth in its service provisions. This is particularly the case when it comes to the privatisation of 'social care', such as residential care facilities, nursing homes and children care services, which the third sector is a central provider of [36–38].

It is commonly argued that the mixed provision of social services have forced TSOs to adapt to contracting and evaluation criteria largely originating from the principles of 'new public management', which, as a result, have forced social service providers to adapt more 'business-

like' features to compete for services [24, 33, 39, 40]. This argument assumes that current social care commissioning criteria emphasizes costs and management practices facilitating performance and quality metrics (such as ISO 9001) that favour 'business aspects' above care aspects [3, 41, 42].

A common way to conceptualise this development is through the concept of *isomorphism*, i.e. the idea that the often homogenous configurations and characteristics of organisations can be explained by how they internalise the similar constraints they are under [43–45]. This theory is based on the notion that organisations are subject to the same environmental features and constraints (by e.g. competing in the same market) and will thus eventually adopt similar characteristics or be eliminated by competition [43]. Whilst isomorphism was originally used to explain the homogeneity of for-profit firms, it is increasingly being used to explain the changing characteristics of TSOs [45–47]. Considering the opening of social services, isomorphism predicts that for-profit, public and third sector organisations will increasingly adopt similar features, which will subsequently result in an increased blurring of the sectors.

However, while this development is often understood as one-directional (i.e. as TSOs and public providers adapting to for-profit characteristics), recent work suggest that this development should be considered to be multi-directional, and that aspects such as the growing emphasis on performance metrics, is the result of a general cultural development of all types of organizations (defined as "constituted as a hybrid of empowered rights and scientific rationality') [48]. This argument is based on the rationale that providers from all sectors are organisations, and have thus changed simultaneously:

> "For instance, a charity responded to some social ill, but a modern nonprofit organization should do so in a way that is accountable, systematic, and effective; a firm

> maximized profits, but a proper one should now also display elements of corporate responsibility; a government could provide public services through a centralized bureaucracy, but a public organization should do so while involving many stakeholders." [48].

Following this line of thought, the argument that sectors are blurring is thus 'overly nonprofit-centric' [48], in that such a position fails to consider more general aspects of the changing organisational culture. While this argument is intuitively strong, it (arguably) fails to consider certain fundamental aspects of sector and how unique characteristics influence the ability to deliver services across ownership. However, there are limited recent and well-developed frameworks to draw on to understand whether and how the mixed provision of social care might influence service and delivery aspects.

Further, one might question the validity of the hybridity assumption underlying the blurring argument. Although the notion of 'hybridity' serves as a useful way of adapting existing concepts to recent observations, it is unclear if the concepts being adapted are best suited to explain organisational change over time. Importantly, hybridity relies on the assumption that institutional and governance logics remain constant over time. In extension, the hybridity notion assumes that organisations entail 'pure forms of activities' that can be traced back to the prototypical typology of sector [127], and that new organisational constellations, such as the social enterprise, should be understood through the original logics associated with the archetypical spheres of the market, state, and civil society [40]. Understanding organisations across sector through the lens of hybridity may thus obstruct the development of new and refined theories on sector activity and distinction. Therefore, I will elaborate on the distinctiveness of the third sector based on a recent and comprehensive analysis on theories on the care market and the third sector. Building

on this analysis, I will develop a revised typology of sector to identify comparative challenges and advantages among for-profit, public, and third sector providers of social services.

'Distinctiveness' of the third sector

Scholars and third sector practitioners have continuously argued that the third sector entails unique characteristics that distinguish it from the public and for-profit sector. Yet, there is little consensus on what specific features can be thought to make the third sector distinct [33, 39]. To approach identifying such distinction, one might ask why the for-profit sector is not the sole private provider of public services, considering that for-profit organisations (FPOs) are often bigger and have more available resources than third sector organisations (TSOs) [49]? Contrary to the belief that more resources lead to more competitive service-deliveries, recent research suggests that TSOs are more effective than for-profit organisations in the context of delivering social care and residential services [38]. This corresponds to the suggested notion that TSOs entail characteristics that make them particularly suitable for public service provision [35]. These characteristics have been identified to include TSOs being better at targeting hard-to-reach population groups, and being more approachable for population groups with specific needs, such as domestic abuse victims or delinquent youths [50].

To conceptualise the distinction between the for-profit, public and third sector sectors the concepts of *legitimacy* and *defining resources* might be used; this is the idea that different types of organisations depend on certain 'core resources'; without which they cannot exist [51–53]. Legitimacy can be defined as the "…*a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions*." [51]; or simply as the "…*conformance with institutional expectations…"* [29]. The idea of legitimizing resources thus revolves around the notion that certain characteristics of an organisation reinforce the legitimacy or the 'license' for that organisation to operate. Obviously, all organisations

depend on certain organisational issues and need to consider, for example, financial or managerial aspects as part of their strategy for sustainability. However, the idea of legitimacy serves as a helpful tool to conceptualise what enables different types of organisations to subsist, which can then inform the construction of strategies regarding how to manage those organisations.

Moore (2000) formalised this idea by developing a model that explains how legitimising factors vary across sectors, which are summarised in the table below.

Table 1: Theoretical distinction between the public, for-profit and third sector (Moore 2000)

	For-Profit Sector	*Public Sector (nonprofit and governmental)*
Normative goal	Enhance shareholder wealth	Achieve social mission
Principal source of revenue	Revenues earned by sale of products and services	Charitable contributions or tax appropriations
Measure of performance	Financial bottom line or increased equity value	Efficiency and effectiveness in achieving mission
Key calculation	Find and exploit distinctive competence of firm by positioning it in product/ service markets	Find better ways to achieve mission

This model rests on the rationale that certain types of organisations are legitimised by specific pursuits ('missions'), which are quintessential for their existence and their 'license to operate'. For example, Moore's model argues that the main pursuit of for-profit sector organisations can be said to revolve around maximising profits, whereas the defining pursuits of the public and third sector is focused on producing 'social value'. Notably, this model does not distinguish between the public and non-profit sector in terms of their mission (social value), which fails to incorporate nuances of how, for example, governments are legitimised by their constituencies, whereas TSOs are not. Whilst it may be argued that the public and third sector are both legitimised by the ability to produce perceived social value, the way this is achieved may vary between the two sectors. First, the ability of the third sector to achieve desirable outputs is often

decided by their funders [54–56], which can vary between individual donors (i.e. the general public) and for-profit or public commissioners, who often impose different guidelines and criteria for effectiveness. The ability of the public sector to create social value is rather decided by the constituencies and public opinion. Further, the revenue of third sector is often constituted by charitable donations and commissioning, while the public sector largely depends on tax revenues to finance its activities.

Moore's model is thus rather flawed in suggesting that there is no distinction between the strategic management of the public and the third sector. However, in terms of explaining the increased third sector delivery of public services, Moore's model is helpful, in that it offers a way to conceptualise why the common pursuit of social value could facilitate reciprocally beneficial partnerships (and thus commissioning) between the third and public sector. Below I will elaborate on Moore's model by adding a few dimensions overlooked in the existing model.

Towards a (clearer) typology of sector

'Sector' is a concept subject to multiple conceptualisations. Table 1 displays a rough distinction of sector, outlining some of the key theoretical features of different types of organisations. However, it is important to note that this rough distinction fails to consider certain nuances and overlaps across different types of providers. Thus, the below table displays theoretical and archetypical features across ownership, which are not empirically validated. As such, the features displayed below are purely conceptual and the validity of these need to be tested in practice.

Table 2: Revised typology of sector (adapted from Moore 2000)

	For-profit sector	**Public sector**	**Third sector**
Definition	Organisations that are privately owned and administered according to the pursuit of profits.	Organisations that are publically owned and administered.	Organisations that are formally structured, privately owned, non-profit distributing, self-governing, and benefiting from voluntary activities [57].

Primary activities	Sale of products and private goods, but increasingly also public goods	Public goods	Public goods, but increasingly also private goods
Main type of income	Profits from sales of products and/or services	Taxation	Charitable contributions and/or commissioning
Purpose/motivation	Maximise and sustain profits	Public sector motivation	Prosocial; altruistic
Stakeholders	Shareholders; customers	Citizens; beneficiaries	Beneficiaries; donors
Management	Flexible and profit-oriented	Inflexible and bureaucratic	Flexible and purpose-oriented

For-profit organisations are often characterised by their pursuit of profit-maximization and shareholder management, whereas public sector providers are often understood as being publically funded and administered by civil servants. Many suggest the voluntary aspect of TSOs, both in terms of labour and monetary donations, as their defining trait [58, 59], in that it separates them from public providers, who largely depend on taxes, and for-profit providers who subsist on profits. However, that is not to say that public sector organisations (PSOs) and TSOs do not benefit from sales of products and services, but simply that their *dependence* on profits differ, in that both public and third sector providers rely on other sources of revenue.

Further, the type of leadership also varies qualitatively among sectors. For example, TSOs are accountable to a board of trustees, whereas for-profit organisations (FPOs) are mainly held accountable by their shareholders. This distinction is substantial in that charity trustees are rarely remunerated monetarily and are, by definition, required to ensure that third sector activities serve the public interest. For-profit shareholders, on the other hand, are often chosen due to their vested interest in the ability of a firm to produce profits.

Moreover, both FPOs and TSOs are more flexible in their organisation and delivery of services compared to public providers, who rely on pre-existing and bureaucratic structures. Yet, it is important to note that bureaucracy has important features that align to the overall purpose of PSOs; for example, it enables adherence to regulation and safeguards against harm and

corruption. However, the motivation to the organisation design can be thought to vary among profit and third sector providers, in that FPOs are driven to organise their service according to their profit pursuit, whereas TSOs will design their operation according to their social purpose.

Unique characteristics influencing different types of service providers

Identifying the unique characteristics across sector is central to understand important aspects of the mixed service provisions of care. Importantly, understanding how and what factors influence providers differently enables an understanding of what specific advantages and disadvantages providers across sectors face. Billis and Glennester (1998) address this in their seminal article by posing that the third sector has a comparative advantage in delivering 'human services', mainly due to their 'ambiguous and hybrid structures', which enable them to, for example, connect to disadvantaged population groups and tackle complex social issues in a more flexible way than public and for-profit providers are capable of. While this analysis stills holds, there are important aspects to add to this framework, which will clarify the specific constraints of different provider types. These are summarised in Table 3 and below.

Table 3: Challenges and advantages in delivering social care across sectors

	For-profit sector	Public sector	Third sector
Main challenges	– Balancing cost with quality – Opportunism – Competing for pro-social employees	– Bureaucracy – Competition with FPOs and TSOs	– Donors and commissioning requirements – Matching social mission with financial support – Competition with other third sector providers
Main advantages	– Financial autonomy – Commissioning and contracting requirements increasingly based on NPM	– Familiarity with rules and regulations – Public sector motivation	– Voluntarism – Prosocial mission and altruism – Community-driven – Eligible for social purpose funding

Challenges across sectors

FPOs are often understood as rational profit maximization entities acting primarily according to the pursuit of profit maximisation. However, to deliver high quality social care, it is essential to have competent and socially motivated staff. Existing literature often suggests that 'mission' is important to attract pro-social individuals and that public and third sector organisations thus have a comparative advantage in recruiting this type of staff [60, 61]. This has been suggested to be one of the main issues for FPOs, who tend to attract profit-oriented individuals, who are more likely to engage in opportunistic behaviour. Following this line of thinking, it is possible to interpret the increasing focus on social efforts, such as corporate social responsibility (CSR), as a strategy to attract more pro-socially motivated workers, following the assumption that 'morally motivated workers' work harder and are less likely to shirk [62, 63]. This idea is consistent with the notion that FPOs are not necessarily becoming more 'socially aware'; rather they are responding to an environment in which having a socially responsible public profile is a necessary condition of being a competitive provider.

It is commonly argued that both public and third sector providers are motivated by 'social mission' [53], which can be interpreted as these providers being driven by prosocial values. In fact, it has been argued that it is the value congruence among them that facilitate the increased third sector provisions of public services [64]. While there is undoubtedly validity to the claim that public and third sector providers are both socially oriented, there are important distinctions in their core values. First, although TSOs are often driven by a social cause or mission, they nevertheless compete for scarce resources against other mission-driven organisations and tend to favour their own work above others, regardless of the potential opportunity costs to the overall social cause of funding a less effective provider (this type of TSOs are referred to as 'warm glow charities' [65]). Further, with the growing reliance on donations and commissioning contracts, TSOs are increasingly dependent on the contracting requirements defined by funding authorities,

which often include performance and evaluation metrics [66]. Thus, to ensure financial survival, TSOs are incentivised to design their service and performance evaluations (such as they exist) such that they are most likely to produce favourable outputs [67, 68]. Yet, it is important to note that this issue is not due to the individual TSOs, as much as it is the result of a flawed donation and commissioning system, which often facilitates incentives to conduct quick and superficial evaluations. Further, the incentive to develop evaluation systems designed to produce positive results may also influence public and for-profit providers; however, given that TSOs often rely entirely on grants and philanthropy, this issue can be thought more important for TSOs.

Organisational mission also influences third sector service delivery. Following the evidence-based practice (EBP) movement, funders are increasingly requiring organisations to deliver programmes that have been validated using rigorous evaluations. In fact, there is a large repository of evidence-based interventions (EBIs) available outlining interventions and programmes that have been associated with positive outcomes [69–71]. While this is not necessarily a negative development, it is important for successful implementation that EBIs can be adapted to suit the needs of the target population defined by their social mission [72]. Considering the central role of 'mission' in TSOs, it follows that providers should carefully evaluate – and be evaluated against – the match of an intervention and its organisational mission before implementation.

PSOs are politically oriented bodies that will inevitably be influenced by the political pursuit to attain the favour of the median-voter [58]. While it is unclear whether this pursuit can convincingly be argued to influence or indeed compromise the operation of social care, it suggests that public providers may prioritise political interests above service quality if presented with the choice. Further, PSOs are conventionally considered to be limited by their bureaucratic features, which on the one hand, may constrain their flexibility in delivering services, but on the

other, ensures that the providers adhere to existing rules and regulations. Last, public providers are completely reliant on government budgets and the quality of their services are thus largely a function thereof.

Advantages across sectors

The main advantage of FPOs is their ability to coordinate in competitive markets and design their allocation of resources accordingly. For example, for-profit providers can allocate their resources without considering the legal restrictions of third and public sector organisations, and can also pursue profit-driven investments. Also, if the 'opening of public services' include allocation principles based on 'new public management' (NPM) (Hood, 1991), the for-profit sector may be at a comparative advantage to PSOs and TSOs when bidding for contracts [39].

As with TSOs, it is often assumed that PSOs attract pro-social workers with public sector motivation (PSM) [73, 74], which can be defined as "an individual's predisposition to respond to motives grounded primarily or uniquely in public institutions" (Perry 1996). This type of individual has been associated with a range of positive traits, such as being driven, responsible, and morally motivated. Recent evidence suggests that 'honest' individuals are more likely to self-select into public services, thus serving as an advantage for these kinds of providers [75]. Public providers may also be at comparative advantage in terms of securing service contracts by knowing the complex bidding system [76]. Further, PSOs may be more familiar with the regulations and best-practice guidelines that social care providers are legally required to adhere to.

Research suggests that the volunteer and community-driven features of TSOs help them recruit hard-to-reach individuals and also allow this type of organisation to engage social issues in a more flexible and communal manner than other providers [58, 77, 78]. That is, the ability of

TSOs to attract motivated employees in communities of need and deprivation is one of the central advantages in their ability to deliver social services. Last, as the for-profit pursuit enables FPOs to pursue funding from investors, third sector providers can apply for grants that have been earmarked for social and non-profit purposes [79].

A revised framework of comparative advantage and disadvantage

All sectors include organisations; yet, these organisations face different types of challenges not only due to their characteristics, but also because of the unique restrictions and opportunities that apply to them. For example, a recent study explored 'sector regret' to find that certain Fair Trade organisations expressed regret in relation to their choice of sector. The regret related to becoming an FPO was often connected with being disqualified from certain grants only available to non-profit providers, whereas the regret of becoming a TSO was related to the flexibility of operation and resource allocation enabled by being a for-profit firm [79]. The comparative advantages and disadvantages are summarised in Table 4.

Table 4: Comparative advantages and disadvantages in delivering social services across sectors

	Employee motivation	**Delivery flexibility**	**Financial flexibility**	**Regulation adherence**
For-profit sector	Disadvantage	Advantage	Advantage	Ambiguous
Public sector	Advantage	Disadvantage	Ambiguous	Advantage
Third sector	Advantage	Advantage	Disadvantage	Ambiguous

Key: *disadvantage:* the characteristics of the provider involve comparative challenges; *advantage:* the characteristics of the provider puts it at comparative advantage; *ambiguous:* the characteristics of the provider is associated with neither a comparative advantage or disadvantage.

Under conditions of traditional competitive markets, FPOs are considered superior to public and third sector providers, in that they have more available resources [80]. In fact, economists consider the growing third sector provision of social services a puzzle, in that this type of provider should, following traditional analyses of financial incentives, be at a comparative disadvantage to FPOs. However, the context of social care services might entail important

distinctions to 'regular' markets, in that care is not only allocated to the highest bidder, but rather constitutes a complex market of need, regulation, subsidies, and self-payment, which vary according to different patient groups. For example, it is important to consider inverse incentives in the mixed delivery of social care, in that it is known that FPOs can – and in the context of hospitals do – cherry-pick their clients according to their treatability and thus design their services to mainly include uncomplicated and inexpensively treated clients. Moreover, a recent study investigating major chains of nursing home providers found that for-profit nursing home chains were generating high revenue, although demonstrating lower quality services than public and third sector providers [81]. This suggests that quality of care may not be the defining factor in the social care market.

By attracting altruistic and pro-socially motivated workers, TSOs should (theoretically) be able to provide care at a higher quality and lower cost than FPOs. This idea aligns to the very conventional (and contested) notion that the third sector is a response to state and market failures, in that they are driven to address issues not satisfactorily tackled by public services. Yet, a more nuanced way of understanding this issue might be by considering public, for-profit, and third sector providers in a synergistic relationship, with TSOs addressing issues not easily addressed through the bureaucratic restrictions of PSOs, and with public authorities providing financial support for both for-profit and third sector services towards population groups not easily or effectively reached by themselves. For example, TSOs may customise their clientele to focus on complicated clients, i.e. those that are not treated by other types of authorities due to risk, costs, or other complications. In fact, recent evidence reported that TSOs working with deprived and/or vulnerable population groups were more likely to receive government funding [82] suggesting that this type of provider were often supported by the state to alleviate deprivation.

The terminology and underlying meaning in relation to sector is subject to different conceptualisations across different fields of research. The above section has attempted to bridge research within economics, public health, political science, and social policy to establish common ground on the understanding of sector in the mixed provision of social services. The main argument is that sector, as defined in this thesis, does matter, especially in the delivery of social services. Although there are nuances not captured in this framework, there are inevitably common external restrictions of FPOs, PSOs, and TSOs that may enable some generalisation across sectors that will matter to the policy context described below.

Evidence-based practice and policy-making

Evidence based practice

In 1969, Donald Campbell published his milestone paper 'Reforms as Experiments' which introduced the Utopian idea of the 'Experimenting Society' (Campbell 1969). The fundamental idea behind the Experimenting Society is constituted by the notion that one should "…*attempt experimental evidence of all program innovations for which such is possible and useful*." (Campbell 1991: 229). That is, interventions in the form of programmes or policies should, to the extent possible, be constructed so that their impact can be measured and evaluated by means of scientific evaluation and critical appraisal.

The 'Experimenting Society' thus focuses on eliminating expert-based decision-making by placing critical appraisal in the centre of evaluation (Donaldson et al. 2009: 3). While expert advice may prove helpful in certain contexts, it should never stand alone in informing or justifying policy decisions, but should rather serve as *"…wise conjectures that need to be tested in implementation*." (Campbell 1991: 229). In turn, the role of the methodologists is central to the 'Experimenting Society', in that they constitute the experts of evaluation. However, Campbell argues that it is not the responsibility of methodologists to advocate "…*what is to be done, but rather to say what has been done*." (ibid 228). That is, the significance of social science primarily relates to testing the effectiveness of programmes, but the actual design of 'innovations' and programmes should be undertaken in consultation with the appropriate stakeholder groups that are "…*responsive to a wider range of interests and experiences than scholars*." (Shadish et al. 1991: 140).

Much has happened since Professor Campbell initially developed the utopian notion of the 'Experimenting Society', and several of its underlying ideas have since taken root in modern evaluation practice. In many ways, these ideas can be argued to constitute the foundation behind

the principles of evidence-based practice (EBP) [85], which are here defined in a healthcare context as "…*the integration of clinical expertise, patient values, and the best evidence into the decision making process for patient care."* [86]. EBP can thus be conceptualised as the interception of expertise, stakeholder values and best available research (see figure below).

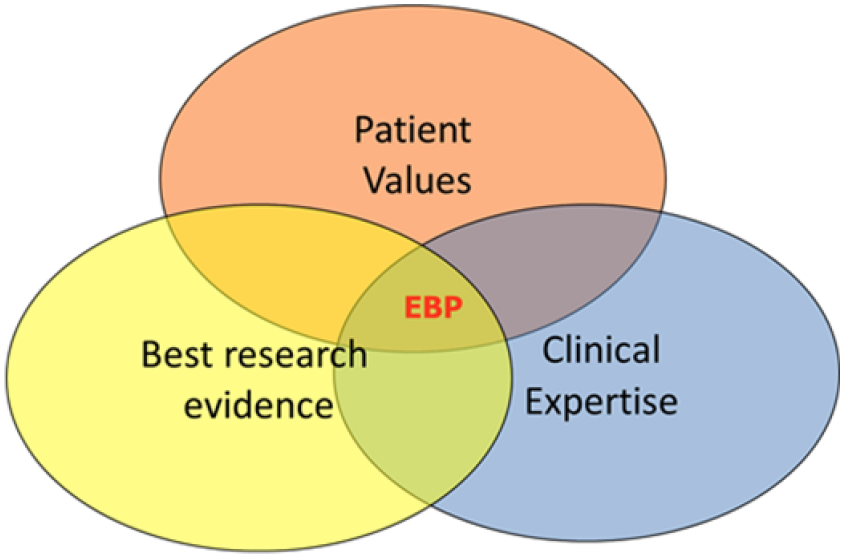


Figure 2: Illustration of EBP

The EBP paradigm is also influenced by the principles of evidence-based medicine [89], which entails the notion of the *hierarchy of evidence*, which is based on the fundamental assumption that certain research evidence can be considered more 'credible' than other types. The hierarchy of evidence traditionally places randomised controlled trials (RCTs) and meta-analyses of these on top, meaning that the aggregation of RCTs constitute the highest possible type of evidence (see figure 3).

Whilst the strict adherence to the hierarchy of evidence is (arguably) losing traction in the EBP [92] movement, the idea that critical appraisal is at the heart of evaluation remains a core

principle in EBP. However, rather than having a fixed and inflexible conceptualisation of the 'best types' of evidence that prioritises the use of RCTs, one ought to consider the *question of interest* before deciding on the most appropriate methodology [85]. That is, different research questions warrant different approaches, and it is thus the specific focus of interest that should define the appropriateness of the applied methodology. This thesis will follow a nuanced definition of 'credible evidence', effectively assuming the relevance of evidence should be decided according to the context of the research and the question of interest.

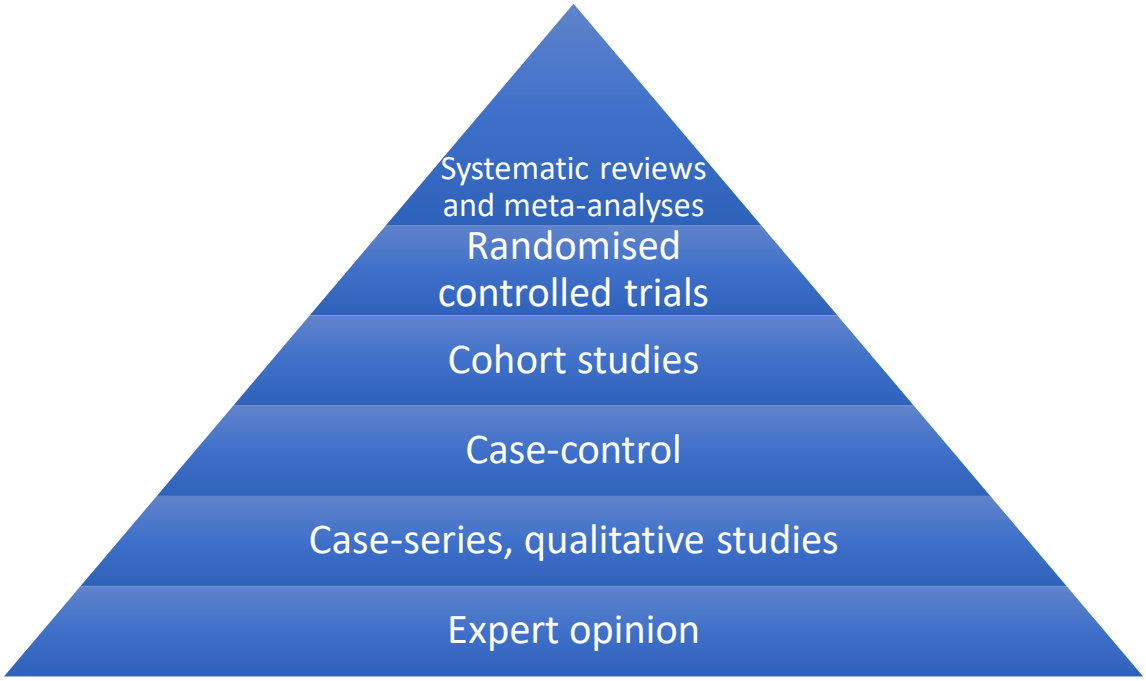


Figure 3: Hierarchy of evidence (Adapted from Greenhalgh 2010)

Evidence-based policy-making

The principles of evidence-based practice are not only advancing in clinical and social work settings [87], but are also becoming increasingly integrated aspect in the policy agenda [88, 89]. While there is mixed evidence on whether current policy-making is actually becoming more evidence-based, as opposed to simply being increasingly utilised as a political discourse tool [89–92], there has been a growing emphasis on 'what works' in the discourse of policy-making. For example, the Obama Administration introduced the 'Evidence-Based Policy Making

Commission Act of 2016', which constituted a public commitment to prioritise the commissioning of activities that are based on sound scientific evidence, and to generally improve the use of data in policy-making [13]. Similarly, other European countries, such as the UK (which will be explored below), have made regulatory advances to ensure that public commissioning is done in line with evidence-based practice [93].

Evidence-based policy making in the UK

UK policy-makers have increasingly adopted a discourse that advocates for the commissioning of public services to be based on the effectiveness of service providers [94]. This focus on service effectiveness is reflected in a range of recent Cabinet Office papers [6, 7, 9, 15], which all emphasise the importance of monitoring of services and 'intelligent commissioning'. For example:

> "Providers of public services from all sectors will need to publish information on performance and user satisfaction." [6]

> "To support better commissioning and innovation in public services, open public services require robust accreditation of what works. Both commissioners and providers need to know which programmes are proven to work." [6]

> "Commissioners need to be enabled to think strategically and take a holistic approach to understanding needs and identifying appropriate outcomes and results. By not taking account of the whole system, decisions can have unintended consequences which cause perverse incentives and poor outcomes." [95]

One might argue that this discourse is a product of the 'Third Way' of government (Alcock 2016), originally introduced by the Labour Government and developed with Anthony Giddens (2003, 2009). The 'Third Way' was introduced in the 1990s and entails the notion that to improve the services of the modern welfare state, one must include a larger range of providers to deliver public services (Alcock 2016). The main rationale behind this argument is that the focus of services should be on 'what works' as opposed to who is providing the specific service [6, 7, 9]. In implementing this 'modernisation of government', bids for public services have gradually become open to more service providers; both within the third and for-profit sector [89, 96].

How this applies to third sector service provisions

In 'opening up' in terms of the public service monopoly, the UK government have specifically called for the participation of the third sector in the delivery of public services [7, 9]. The commitment to have the third sector participate in the delivery of public services is, for example, reflected in the relatively recent White Paper 'Building a Stronger Civil Society' (2010), in which the Government called for the increased participation of the third sector in public services.

> "The Government is committed to ensuring that charities, social enterprises and cooperatives have a much greater role in the running of public services." [9]

> "[The government will] Ensure that CSOs [civil society organisations] have a greater role and more opportunities in delivering public services by opening up new markets in accordance with wider public service reform measures and reforming the commissioning environment in existing markets."[8]

In so doing, the 2010 White Paper specifically requests increased third sector participation in the areas of (1) Health and Social Care, (2) The Criminal Justice System, and (3) Welfare to Work;

areas that all involve social interventions and vulnerable population groups. This encouragement of third sector participation in public services is in line with the idea that the state, market and third sector is becoming increasingly embedded in the alleviation of social issues [24, 29].

Whilst this development is often explained by the paradigm of the 'Third Way', others have argued it is simply a method of making the third sector more 'governable' [34, 96]. That is, by enabling the third sector to partake in the market of public services, the third sector subsequently becomes more dependent on the requirements and regulations set by public commissioners through contracting; which most third sector commissioning is based on [96, 97]. To this end, it has been argued that the contracting objectives of TSOs often take a narrow perspective on performance by focusing on business related aspects (e.g. pay-for-performance schemes), rather than aspects related to the capacity to reach and work with vulnerable groups [96, 98]. This idea is consistent with the claim that TSOs experience an increasing pressure to evaluate their activities in legitimising their activities following a 'new public management' line of thinking [17, 66, 99, 100].

Evaluation and implementation among third sector providers of public services

In 2012 the UK Coalition government initiated the Social Value Act, which became effective from 2013 and set out to increase the knowledge of the *social value* produced by services delivered or commissioned by the public sector [102]. In so doing, the act "…*require[s] public authorities to have regard to economic, social and environmental well-being in connection with public services contracts; and for connected purposes*." [103]. However, a 2015 report on the impact of the Social Value Act [15] found that there is little consistency in terms of the conceptualisation and measurement of social value and impact, which made it challenging to enforce organisations to produce reports evidencing their impact. This suggests that while there may be *political will* to require social impact

evalaution in the delivery of public services, there is an absence of commonly agreed and operationable reporting frameworks and guidelines that can be enforced through regulation.

The failure to specify terms like 'social impact' and 'performance' forms a situation, in which UK charities are required to demonstrate and elaborate on how they achieve their 'public benefit', but without being subject to a clear definition of what that actually means [104], which (presumably) makes adhering to such regulation rather challenging. Indeed, a recent report by the Charity Commission [105] which surveyed a random sample of yearly reports (n=107) found that more than half of that sample (54%) did not report (according to the CC) satisfactorily on how they achieve their public and social benefit. Whilst the overall sample of the report is rather small, it seems to suggest that the current regulation is not adhered to by many TSOs. However, considering the finding of the Social Value Report, this report may also reflect the fact that there are no commonly agreed-on evaluation criteria in place to enforce consistent public benefit reporting. Also, considering that the CC needed to conduct a study to establish whether TSOs adhere to the public benefit requirement, seems to suggest that there are no monitoring practices in place to make sure that TSOs are actually reporting their public benefit.

When it comes to assessing the impact of third sector service deliveries, and whether the substitution of third sector delivery of public services leads to lower costs and more effective service deliveries, the evidence is scarce [10, 36]. In fact, most studies on the subject simply conclude that too little is known about the use of evidence by TSOs, and that more representative studies are needed to map their service delivery [10, 32, 36]. This lack of quantitative data on the third sector service delivery was echoed in an open letter by Andrew Dilnot to the Charity Commission:

> "We see a shortage of regular and timely data in this area and we think it likely that this will inhibit the monitoring of the beneficial impact of the voluntary sector and its contribution to the Big Society and other Government initiatives." [106]

This lack of 'regular and timely' data is contrary to the above cited claims of the UK government to pursue more transparency on the delivery of publicly commissioned services. This absence of aggregate data seems contrary to the fact that TSOs are reporting to be under increasing pressure to evaluate their work to compete for the commissioning of services [66, 107, 108]. However, as demonstrated by the regulatory framework of UK TSOs, there is currently very limited guidance on *how* the third sector should evidence its activities.

Whilst it is normally argued that the increased pressure for TSOs to evaluate is a function of the pursuit of *public accountability* of the public sector, who often fund those services [104], the tendency that commissioning requirements are often built around business templates [29, 39, 41, 54] suggests that the pressure to evaluate may be a result of isomorphism. Having organisations adapt their characteristics to compete in a changing market is not necessarily a negative development, in that market disruption can lead to improvements through innovations. However, studies that investigate the evaluation capacity of TSOs find that such organisations tend to customise their evaluation procedure not to improve practice, but rather to satisfy funding bodies and to meet the increased pressure of performance measurement [16, 17, 109–112].

This is highly problematic, in that the motivation behind evaluation activities should be to ensure that the provided services are safe and effective; especially considering that most interventions have the potential to do harm to those they target [19, 22, 113, 114]. Though interventions rarely *intend* to do harm, evidence has continuously shown that interventions, which are based on

sound *programme theories*, do not always produce the anticipated results in practice (e.g. Petrosino et al. 2003, McCord 2003, Lilienfield 2007, Pignotti and Thyer 2009, Thyer and Pignotti 2011). However, the risk of adverse or iatrogenic effects is arguably always present, in effect making a pursuit towards harm-free interventions naïve as well as unrealistic. Rather, the inevitable presence of risk is exactly why adhering to 'best practice' and sound critical appraisal is central in ensuring the safety of service-users; especially in the context of vulnerable populations. Therefore, a commissioning market built around superficial evaluation measures, which effectively incentivises poor evaluation poses a significant risk to service-users. In fact, the risk of poor evaluation may constitute a higher risk than no evaluation, given that activities evaluated using poor practice may be allowed to continue, despite being harmful or ineffective.

Policy assumptions underlying the opening of public services

In opening up public service provisions to market bidding, it is central to consider the framework under which providers are incentivised to operate. As addressed previously, providers from different sectors will inevitably be subject to different conditions that influence their ability to deliver services effectively and in accordance with best practice. Thus, the successful opening of services relies on a number of assumptions. These are summarised below.

(1) Third sector providers are capable of evaluating their services reliably.
(2) Third sector providers are capable of implementing their services according to best practice.
(3) The mixed provision of care will improve service quality rather than compromise it.

This thesis will address each of these assumptions using best-practice guidance and a range of rigorous research methods. However, it is important to note that these assumptions are not exhaustive; yet, they represent a number of important issues relevant for the third sector delivery of social care services. Each assumption is developed below.

Papers 1 & 2: What issues do TSOs experience when evaluating and implementing their services?

To understand how the third sector delivery of social and health services might be *improved*, one must first grasp what needs and issues TSOs face as practitioners. Organisational needs and issues have been operationalised in the context of *evaluation* and *implementation*; both central features for the adherence to EBP, and often a blind-spot for TSOs. In terms of investigating these issues, the thesis employs the systematic review methodology following best practice reporting guidelines, which is widely considered the gold standard for the identification,

appraisal, and synthesis of existing research. The focus of these systematic reviews can be summarised by the following questions:

- What barriers and facilitators do third sector practitioners identify in relation to evaluating the services their organisations provide?
- What barriers and facilitators influence the implementation process of third sector organisations delivering evidence-based interventions and programmes?

The findings of the systematic reviews provide an empirical background in terms of the factors that facilitate and obstruct TSOs to deliver and evaluate services. This background constitutes an important step in terms of addressing how future policy may facilitate better regulatory and commissioning practice.

Methods

Systematic reviews are traditionally designed to aggregate, estimate and evaluate the evidence of effectiveness of interventions [115–117]. There are several established guidelines to ensure rigorous extraction and synthesis of intervention studies, such as the PRISMA guidelines [118]. However, to date no consensus-based reporting and data-synthesis guidelines have been modified to systematic reviews that investigate factors that influence organisational ability. That is, there is no 'gold-standard' in relation to the conduct of systematic reviews on barriers and facilitators. As a result, systematic reviews focusing on barriers and facilitators often involve a 'narrative synthesis' or 'vote-counting' approach, both of which are considered the weakest type of data-synthesis [115]. To mitigate this issue, care has been taken to refer to the number of participants in each of the studies reviewed so as to give readers balance of the strength of opinion. It should, however, be remembered that studies testing these questions are largely qualitative and thus have small Ns relative to quantitative effectiveness trials.

Whilst there is little consensus on the appropriate conduct of systematic reviews that do not focus on effectiveness of interventions, reviews of barriers and facilitators are frequently published in high impact peer-reviewed academic journals, despite not being able to fully adhere to traditional methods of data synthesis [32, 91, 119, 120]. For example, systematic reviews on barriers and facilitators often include a variety of research designs such as interviews, case studies, focus groups, observations and survey studies, in which no 'estimate of effect' can be given. Mixed methods synthesis is therefore often the most desirable method in this type of review; especially as reviews of barriers and facilitators often aim to aggregate and describe the full body of evidence.

To synthesise the findings of the included studies, this thesis constructed thematic categories to aggregate the identified factors under overarching themes [91], thus allowing for a framework that captures the overall themes in a way that accounts for the arbitrary difference between identified factors that are similar in nature. Further, this thesis quality appraised all included studies and investigated whether the findings of the review varied across studies of different quality. Specifically, the thesis applied the recent AXIS checklist for cross-sectional studies [121] and an adapted version of the checklists developed by the Joanna Briggs Institute and CASP to evaluate the quality of interview, observation and case study research. Additionally, the findings of the studies were organised into a table illustrating how each study contributed to the construction of the identified themes [122].

Paper 3: Quality variation among for-profit, public, and third sector social care providers

Following the investigation of the evaluation and implementation issues of TSOs, the thesis proceeds to empirically explore whether the third sector varies from public and for-profit sector organisations in terms of the quality ratings provided by the Care Inspectorate; the independent regulator of social care services in Scotland. The Care Inspectorate serves as the regulator of care

services (e.g. child care, nursing homes, etc.) and currently (June 2017) includes data on 13,522 organisations. Of these organisations, 10.5% work with nursing homes, whereas the rest are split among a list of types of services, such as day care for children and various types of support services.

Researchers, policy-makers and practitioners continue to argue that the third sector entails distinct, and often favourable, features making them unique providers of services [29, 32, 53]. While it is known that the provision of many social services in the UK is delivered through a quasi-market of public, for-profit and third sector providers, little is known about whether there are systematic differences in service quality according to different provider types. It is often hypothesised that for-profit organisations are legitimised by their ability to maximise profits, effectively suggesting that their activities can be explained considering by this pursuit, and, conversely, that third and public sector organisations are primarily legitimised by their ability to achieve *social value,* suggesting that their incentives to deliver services are predominantly motivated by pro-social pursuits. However, there is little empirical evidence that test what these theoretical differences between sectors mean in practice and how and whether they translate into, for example, quality of services.

To add to this body of research, this paper tests whether sector does indeed serve as a meaningful predictor of the quality of care services using data from the Care Inspectorate in Scotland. The question can be summarised thus:

- Does the quality of social care service vary across for-profit, public and third sector organisations as assessed by the quality and risk domains of the Care Inspectorate in Scotland?

Testing the quality of the third sector delivery in social care settings allows for testing if sector predicts variation in service quality. Testing the 'effect' of sector will inform theory on the typology of sector, which traditionally suggests that the distinctiveness of the sectors influences the quality of service deliveries; an assumption that remains largely untested.

Methods

The analysis was conducted using multiple generalised ordered logistic regression models. The main outcomes are the quality and risk domains collected and assessed by the Care Inspectorate, which include the following areas: (1) Quality of care and support, (2) Quality of Management and Leadership, (3) Quality of Environment, (4) Quality of Staffing, and, (5) Risk of services. To test the robustness of the outcomes collected and assessed by the Care Inspectorate, the analysis also investigated whether ownership explains variation in user-reported outcomes, such as complaints. Considering recent criticisms of the relativistic features of odds ratios [123], the results will be reported as predicted probabilities, which are considered a more robust and intuitively readable effect size measure [124, 125].

The final models will include all relevant control variables allowed for in the dataset to ensure that the effect of sector on service quality is not confounded by, for example, the type of services, target groups of care organisations, level of risk of the service activities, or whether an organisation has been subject to any forced requirements as result of an inspection. The results will be subject to comprehensive robustness analyses and reported following best practice reporting guidelines [126].

Conclusion

From a regulation perspective, the large variety of activity within the third sector makes it difficult to develop policies that are appropriate to the full range of TSOs. While policy makers and academics continue to emphasise the growing importance of the third sector as a social service provider [6, 10–12], TSOs are subject to little guidance on how to evaluate and implement their services following EBP. Further, there is limited evidence on the effectiveness of the mixed provision of care services in the UK. This lack of clarity on how the third sector should evaluate is problematic and essentially comes at the expense of the beneficiaries; most of which include vulnerable population groups.

The findings of this thesis have important practical and academic implications, which will be described in full. Perhaps of most significance, the findings constitute an important resource to inform strategies on how the third sector delivery of health and social services can be improved on a national scale and perhaps beyond. Specifically, the thesis provides substantial contributions to the understanding of the (1) evaluation practices, and (2) delivery needs of TSOs, and, (3) the distinctiveness of the third sector as a service provider compared to for-profit and public providers. In sum, these findings contribute to an up-to-date formulation of strategies regarding ways in which the third sector may better incorporate the principles of evidence-based practice in delivering social services.

Bibliography

1. Le Grand J. Quasi-Market versus State Provision of Public Services : Some Ethical Considerations. Public Reason. 2011;3:80–9.
2. Grand J Le, Le Grand J. Quasi-Markets and Social Policy. Econ J. 1991;101 September:1256–67. doi:10.2307/2234441.
3. Atkinson C, Crozier S, Lucas R. Workforce Policy and Care Quality in English Long-term Elder Care. Public Perform Manag Rev. 2018. doi:10.1080/15309576.2018.1473784.
4. Cunningham I, James P. The outsourcing of social care in Britain: What does it mean for voluntary sector workers? Work Employ Soc. 2009;23:363–75. doi:10.1177/0950017009102863.
5. Hardill I, Dwyer P. Delivering public services in the mixed economy of welfare: Perspectives from the voluntary and community sector in rural England. J Soc Policy. 2011;40:157–72.

6. Cabinet Office. Open Public Services: White Paper. 2011. http://www.cabinetoffice.gov.uk/sites/default/files/resources/open-public-services-white-paper.pdf%5Cnhttp://www.openpublicservices.cabinetoffice.gov.uk/.
7. National Council for Voluntary Organisations (NCVO). Green Paper: Modernising Commissioning: Increasing the Role of Charities, Social Enterprises, Mutuals and Cooperatives in Public Service Delivery.. 2010.
8. Government H. The Compact The Coalition Government and civil society organisations working effectively in partnership for the benefit of communities and citizens in England. 2010. http://webarchive.nationalarchives.gov.uk/20110314111751/http://www.thecompact.org.uk/files/143051/FileName/TheCompact.pdf. Accessed 19 May 2017.
9. Office for Civil Society. Building a Stronger Civil Society. 2010.
10. Macmillan R. The third sector delivering public services: an evidence review. TSRC Working Paper 20 (Birmingham, University of Birmingham). 2010. https://www.birmingham.ac.uk/generic/tsrc/research/service-delivery/wp-20-public-services-evidence-review.aspx. Accessed 7 Apr 2016.
11. Hogg E, Baines S. Changing Responsibilities and Roles of the Voluntary and Community Sector in the Welfare Mix: A Review. Soc Policy Soc. 2011;10:341–52. doi:10.1017/S1474746411000078.
12. Kelly J. Reforming public services in the UK: Bringing in the third sector. Public Adm. 2007;85:1003–22. doi:10.1111/j.1467-9299.2007.00680.x.
13. Commission on Evidence-Based Policymaking. About CEP. 2017. https://www.cep.gov/about.html. Accessed 30 Aug 2017.
14. Charity Commission. Charity Reporting and Accounting : The essentials. Account Report. 2013; January:1–28.
15. Cabinet Office. Social Value Act review. 2015.
16. Ellis J. Monitoring and evaluation in the third sector: meeting accountability and learning needs. 2009. p. 1–9. doi:ISBN 978-0-9555404-9.
17. Ógáin EN, Lumley T, Pritchard D. Making an Impact: Impact measurement among charities and social enterprises. 2012; October:1–73. http://www.thinknpc.org/publications/making-an-impact/. Accessed 10 Aug 2016.
18. Chalmers I. If evidence-informed policy works in practice, does it matter if it doesn't work in theory? Evid Policy. 2005;1:227–42.
19. Chalmers I. Trying to do more Good than Harm in Policy and Practice: The Role of Rigorous, Transparent, Up-to-Date Evaluations. Ann Am Acad Pol Soc Sci. 2003;589:22–40. doi:10.1177/0002716203254762.
20. Lilienfeld SO. Psychological Treatments That Cause Harm. Perspect Psychol Sci. 2007;2:53–70. doi:10.1111/j.1745-6916.2007.00029.x.
21. Welsh BC, Rocque M. When crime prevention harms: a review of systematic reviews. J Exp Criminol. 2014;:1–22.
22. McCord J. Cures That Harm: Unanticipated Outcomes of Crime Prevention Programs. Ann Am Acad Pol Soc Sci. 2003;587:16–30.
23. Petrosino A, Turpin-Petrosino C, Buehler J. Scared Straight and Other Juvenile Awareness Programs for Preventing Juvenile Delinquency: A Systematic Review of the Randomized Experimental Evidence. Ann Am Acad Pol Soc Sci. 2003;589:41–62.
24. Alcock P. A strategic unity: defining the third sector in the UK. Volunt Sect Rev. 2010;1:5–24.
25. Evers A, Laville J-L. Defining the third sector in Europe. Edward Elgar Publishing; 2004.
26. Salamon LM. Defining the Nonprofit Sector: The United States. Work Pap Johns Hopkins Comp Nonprofit Sect Proj. 1996;:1–14.
27. Defourny J, Kirsten Grønbjerg B, Lucas Meijs B, Marthe Nyssens B, Naoto Yamauchi B, Grønbjerg K, et al. Voluntas Symposium: Comments on Salamon and Sokolowski's Re-

conceptualization of the Third Sector. VOLUNTAS. 2016;27:1546–61. doi:10.1007/s11266-016-9743-y.
28. Salamon LM, Anheier HK, List R, Toepler S, Sokolowski SW. Civil Society in Comparative Perspective. 1999.
29. Anheier HK. Nonprofit Organizations: Theory, management, policy. Routledge; 2005. http://books.google.com/books?id=s_uzg-KoVyIC&pgis=1.
30. Morris S. Defining the non-profit sector: Some lessons from history. 2000. http://eprints.lse.ac.uk/29032/1/cswp3.pdf. Accessed 29 Mar 2017.
31. Kendall J, Knapp M. Measuring the Performance of Voluntary Organizations. Public Manag Rev. 2000;2:105–32. doi:10.1080/14719030000000006.
32. Hardwick R, Anderson R, Cooper C. How do third sector organisations use research and other knowledge? A systematic scoping review. Implement Sci. 2015;10:84. doi:10.1186/s13012-015-0265-6.
33. Billis D. Towards a theory of hybrid organizations Book section Hybrid Organizations and the Third Sector Chapter 3 Towards a theory of hybrid organizations. Billis Palgrave Macmillan. 2010;:46–69. http://eprints.lse.ac.uk/55260/. Accessed 26 Jan 2017.
34. Kramer RM. A Third Sector in the Third Millennium? Int J Volunt Nonprofit Organ. 2000;11.
35. Rees, J. Mullins, D. The third sector delivering public services: Developments, Innovations and Challenges Policy Press; 2017.
36. Dickinson H, Allen K, Alcock P, Macmillan R, Glasby J. The Role of the Third Sector in Delivering Social Care. NIHR School for Social Care Research. 2012. doi:978-0-85328-443-7.
37. Stolt R, Blomqvist P, Winblad U. Privatization of social services: Quality differences in Swedish elderly care. Soc Sci Med. 2011;72:560–7. doi:10.1016/j.socscimed.2010.11.012.
38. Barron DN, West E. The quasi-market for adult residential care in the UK: Do for-profit, not-for-profit or public sector residential care and nursing homes provide better quality care? Soc Sci Med. 2017;179:137–46. doi:10.1016/j.socscimed.2017.02.037.
39. Dart R. Being "Business-Like" in a Nonprofit Organization: A Grounded and Inductive Typology. Nonprofit Volunt Sect Q. 2004;33:290–310.
40. Pestoff V. Hybridity, Coproduction, and Third Sector Social Services in Europe. Am Behav Sci. 2014;58:1–13. doi:10.1177/0002764214534670.
41. Paton R. Managing and Measuring Social Enterprises. 2003. doi:10.4135/9781446221327.
42. White GRT, Samson P, Rowland-Jones R, Thomas AJ. The implementation of a quality management system in the not-for-profit sector. TQM J. 2009;21:273–83. doi:10.1108/17542730910953040.
43. DiMaggio PJ, Powell WW. The Iron Cage revisited: institutional isomorphism and collective rationality in organizational fields. Am Sociol Rev. 1983;48:147–60.
44. DiMaggio PJ, Anheier HK. The Sociology of Nonprofit Organizations and Sectors. Annu Rev Sociol. 1990;16:137–59. doi:10.1146/annurev.so.16.080190.001033.
45. Mason C. Isomorphism, Social Enterprise and the Pressure to Maximise Social Benefit. J Soc Entrep. 2012;3:74–95. doi:10.1080/19420676.2012.665382.
46. Leiter J. Structural Isomorphism in Australian Non-profit Organizations. 2005. http://eprints.qut.edu.au/4460/1/4460.pdf. Accessed 10 May 2017.
47. Verbruggen S, Christiaens J, Milis K, Henzen MC. Organizations' Compliance With Reporting Standards? Can Resource Dependence and Coercive Isomorphism Explain Nonprofit. 2010. doi:10.1177/0899764009355061.
48. Bromley P, Meyer JW. "They Are All Organizations": The Cultural Roots of Blurring Between the Nonprofit, Business, and Government Sectors. Adm Soc. 2017;49:939–66. doi:10.1177/0095399714548268.
49. Rose-ackerman S. Altruism , Nonprofits , and Economic Theory. 1996;34:701–28.
50. Minzner A, Klerman JA, Markovitz CE, Fink B. The Impact of Capacity- Building Programs

on Nonprofits: A Random Assignment Evaluation. Nonprofit Volunt Sect Q. 2014;43:547–69.
51. Suchman MC. Managing Legitimacy: Strategic and Institutional Approaches. Source Acad Manag Rev. 1995;20:571–610. http://www.jstor.org/stable/258788. Accessed 23 Apr 2017.
52. Dart R. The Legitimacy of Social Enterprise. NONPROFIT Manag Leadersh. 2004;14.
53. Moore MH. Managing for Value: Organizational Strategy in for-Profit, Nonprofit, and Governmental Organizations. Nonprofit Volunt Sect Q. 2000;29:183–204. doi:10.1177/089976400773746391.
54. Barman E. What is the Bottom Line for Nonprofit Organizations? A History of Measurement in the British Voluntary Sector. Voluntas. 2007;18:101–15.
55. Barman E, Macindoe H. Institutional Pressures and Organizational Capacity: The Case of Outcome Measurement. Sociol Forum. 2012;27:70–93.
56. Moxham C, Boaden R. The impact of performance measurement in the voluntary sector. Int J Oper Prod Manag. 2007;27:826–45. doi:10.1108/01443570710763796.
57. Salamon LM, Anheier HK. In search of the non-profit sector. I: The question of definitions. Voluntas. 1992;3:125–51. doi:10.1007/BF01397770.
58. Billis D, Glennester H. Human Services and the Voluntary Sector: Towards a Theory of Comparative Advantage. J Soc Policy. 1998. https://www.cambridge.org/core/terms. Accessed 21 Nov 2018.
59. Witesman EM. An Institutional Theory of the Nonprofit: Toll Goods and Voluntary Action. 2016;45:97–115. doi:10.1177/0899764016651729.
60. Besley T, Ghatak M. Competition and Incentives with Motivated Agents. Am Econ Rev. 2005;95:616–36. https://ezproxy-prd.bodleian.ox.ac.uk:4228/stable/pdf/4132732.pdf?refreqid=excelsior%3Ad988fc7c95d1718e7548468046b821e8. Accessed 3 Feb 2019.
61. Barigozzi F, Burani N. Competition and screening with motivated health professionals. 2016. doi:10.1016/j.jhealeco.2016.06.003.
62. Brekke KA, Nyborg K. Attracting responsible employees: Green production as labor market screening. Resour Energy Econ. 2008;30:509–26. doi:10.1016/j.reseneeco.2008.05.001.
63. Nyborg K, Brekke KA. Moral Hazard and Moral Motivation: Corporate Social Responsibility as Labor Market Screening. 2005. doi:10.2139/ssrn.645741.
64. Miller-Stevens K, Taylor JA, Morris JC, Miller-Stevens K, Morris JC. Are We Really on the Same Page? An Empirical Examination of Value Congruence Between Public Sector and Nonprofit Sector Managers. Volunt Int J Volunt Nonprofit Organ. 2014. doi:10.1007/s11266-014-9514-6.
65. Scharf K. Impure prosocial motivation in charity provision: Warm-glow charities and implications for public funding. J Public Econ. 2014;114:50–7. doi:10.1016/j.jpubeco.2013.10.002.
66. Barman E. What is the bottom line for nonprofit organizations? A history of measurement in the british voluntary sector. Voluntas. 2007;18:101–15. doi:10.1007/s11266-007-9039-3.
67. Bach-Mortensen AM, Montgomery P. What are the barriers and facilitators for third sector organisations (non-profits) to evaluate their services? A systematic review. Syst Rev. 2018;7:13. doi:10.1186/s13643-018-0681-1.
68. Pritchett L. It pays to be ignorant: A simple political economy of rigorous program evaluation. J Policy Reform. 2002;5:251–69. doi:10.1080/1384128032000096832.
69. Blueprints for Healthy Child Development. http://www.blueprintsprograms.com/.
70. CDC - Evidence-Based Practices - Health System Resource - STLT Gateway. https://www.cdc.gov/publichealthgateway/program/resources/evidence.html. Accessed 7 Feb 2019.
71. What Works? 2014. www.nice.org.uk. Accessed 29 May 2017.
72. Bach-Mortensen AM, Lange BCL, Montgomery P. Barriers and facilitators to implementing evidence-based interventions among third sector organisations: a systematic review. Implement

Sci. 2018;13:103. doi:10.1186/s13012-018-0789-7.
73. Wright BE. Public service and motivation: Does mission matter? Public Adm Rev. 2007;67:54–64.
74. Steen T. Public sector motivation: is there something to learn from the study of volunteerism? 2005. http://soc.kuleuven.be/io/egpa/HRM/bern/Steen.PDF. Accessed 10 May 2017.
75. Olsen AL, Hjorth F, Harmon N, Barfort S. Behavioral Dishonesty in the Public Sector. J Public Adm Res Theory. 2018. doi:10.1093/jopart/muy058.
76. Damm C. The third sector delivering employment services : an evidence review. 2012. doi:10.1126/science.aaa5099.
77. Pandey S, Peng S, Pandey SK. Is There a Nonprofit Advantage? Examining the Impact of Institutional Context on Individual–Organizational Value Congruence 585. Public Adm Rev. 2015;75:585–96. doi:10.1111/puar.12357.
78. Mayhew F. Human Service Delivery in a Multi-Tier System: the Subtleties of Collaboration Among Partners. J Health Hum Serv Adm. 2012. https://about.jstor.org/terms. Accessed 5 Feb 2019.
79. Witesman EM, Child C, Breck Wightman G. Sector Choice and Sector Regret. Nonprofit Volunt Sect Q. 2019. doi:10.1177/0899764018824668.
80. Ghatak M, Mueller H. Thanks for nothing? Not-for-profits and motivated agents. J Public Econ. 2010;95:94–105. doi:10.1016/j.jpubeco.2010.09.003.
81. Harrington C, Jacobsen FF, Panos J, Pollock A, Sutaria S, Szebehely M. Marketization in Long-Term Care: A Cross-Country Comparison of Large For-Profit Nursing Home Chains. Heal Serv Insights. 2017.
82. Clifford D, Geyne-Rahme F, Mohan J. Variations between organisations and localities in government funding of third-sector activity: Evidence from the national survey of third-sector organisations in England. Urban Stud. 2013;50:959–76. doi:10.1177/0042098012458550.
83. Mohan J, Bennett MR. Community-level impacts of the third sector: Does the local distribution of voluntary organizations influence the likelihood of volunteering? Econ Sp. 2019. doi:10.1177/0308518X19831703.
84. Clifford D. Neighborhood Context and Enduring Differences in the Density of Charitable Organizations: Reinforcing Dynamics of Foundation and Dissolution. Am J Sociol. 2018;123:1535–600.
85. Donaldson SI, Christie CA, Mark MM. What Counts as Credible Evidence in Applied Research and Evaluation Practice? Evaluation. 2008;6:288. doi:10.4135/9781412995634.
86. Sackett D. Evidence-based medicine : how to practice and teach EBM. Churchill Livingstone; 2000. doi:10.1016/S1031-170X(97)80036-0.
87. Montgomery P, Grant S, Hopewell S, Macdonald G, Moher D, Michie S, et al. Protocol for CONSORT-SPI: an extension for social and psychological interventions. Implement Sci. 2013;8:99. doi:10.1186/1748-5908-8-99.
88. Davies H. Discussion Paper 2 Evidence-based policy and practice: moving from rhetoric to reality. 2002.
89. Nutley S, Davies H, Smith P. What works?: Evidence-based policy and practice in public services. 2000.
90. Smith KE, Stewart E. 'Black magic' and 'gold dust': the epistemic and political uses of evidence tools in public health policy making. 2015;11:415–37.
91. Oliver K, Innvar S, Lorenc T, Woodman J, Thomas J. A systematic review of barriers to and facilitators of the use of evidence by policymakers.(Research article). BMC Health Serv Res. 2014;14:2.
92. Oliver K, Lorenc T, Innvær S. New directions in evidence-based policy research: a critical analysis of the literature. Heal Res Policy Syst. 2014;12:34.
93. Nutley S, Morton S, Jung T, Boaz A. Evidence and policy in 6 European countries: diverse

appraoches and common challenges. Evid Policy. 2010;6:131–44.
94. Davies H, Nutley S. Member of the ESRC Network for Evidence-based Policy and Practice Evidence-based policy and practice: moving from rhetoric to reality. 2002. http://www.ruru.ac.uk/pdf/Rhetoric to reality NF.pdf. Accessed 17 May 2017.
95. National Council for Voluntary Organisations. Modernising Commissioning: Increasing the Role of Charities, Social Enterprises, Mutuals and Cooperatives in Public Service Delivery. 2010.
96. Carmel E, Harlock JE. Instituting the "third sector" as a governable terrain: Partnership, procurement and performance in the UK. Policy Polit. 2008;36:155–71.
97. NCVO. NCVO UK Civil Society Almanac. 2019. https://data.ncvo.org.uk/. Accessed 5 Jan 2017.
98. Cairns B, Harris M, Young P. Building the Capacity of the Voluntary Nonprofit Sector: Challenges of Theory and Practice. Intl J Public Adm Intl J Public Adm Intl J Public Adm. 2005;28:869–85.
99. Arvidson M. Evidence and transparency in the open public services reform: perspectives for the third sector. TSRC Working Paper 117 (Birmingham, University of Birmingham). 2014. https://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-117.pdf. Accessed 6 Aug 2016.
100. Ellis J, Gregory T. Developing monitoring and evaluation in the third sector: Research report. 2008. http://www.poolecvs.org.uk/documents/cesresearchreportaccountabilityandlearning320328.pdf . Accessed 19 Oct 2016.
101. Cairns, B., Harris, M., Hutchison, R., & Tricker M. Improving performance? The adoption and implementation of quality systems in UK nonprofits. , 16(2), 135-151. Nonprofit Manag Leadersh. 2005;16.
102. Cabinet Office. Social Value Act: information and resources - GOV.UK. 2016. https://www.gov.uk/government/publications/social-value-act-information-and-resources/social-value-act-information-and-resources. Accessed 30 Aug 2017.
103. Cabinet Office. Public Services (Social Value) Act 2012. Queen's Printer of Acts of Parliament; 2012. http://www.legislation.gov.uk/ukpga/2012/3/enacted. Accessed 30 Aug 2017.
104. Moxham C, Boaden R. The impact of performance measurement in the voluntary sector - identification of contextual and processual factors. Int J Oper Prod Manag. 2007;27:826–45. doi:10.1108/01443570710763796.
105. Charity Commission. Accounts monitoring review Telling your story well: public benefit reporting by charities. 2017. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/609953/AMR_Telling_your_story_well_public_benefit_reporting_by_charities.pdf. Accessed 10 May 2017.
106. Dilnot A. Letter from Andrew Dilnot to Nick Hurd. 2012.
107. Mitchell GE, Berlan D. Evaluation and Evaluative Rigor in the Nonprofit Sector. Nonprofit Manag Leadersh. 2016;27:237–50.
108. Kendall J, Knapp M. Measuring the Performance of Voluntary Organizations. Public Manag Rev. 2000;2:105–32. doi:10.1080/14719030000000006.
109. Carman JG, Fredericks KA. Nonprofits and Evaluation: Empirical Evidence From the Field. In: Nonprofits and evaluation. New Directions for Evaluation. 2008. p. 51–71.
110. Carman JG. Evaluation practice among community-based organizations: Research into the reality. Am J Eval. 2007;28:60–75. doi:10.1177/1098214006296245.
111. Jamieson M. Author's note. Educ Prim Care. 2011;22:123.
112. Millar R, Hall K. Social Return on Investment (SROI) and Performance Measurement. Public Manag Rev. 2012;9037 July 2015:1–19. doi:10.1080/14719037.2012.698857.
113. Dishion TJ, Mccord J, Poulin F. When Interventions Harm: Peer Groups and Problem Behavior. Am Psychol . 1999;54:755–64. doi:10.1037/0003-066X.54.9.755.

114. Petrosino A, Turpin-Petrosino C, Buehler J. Scared Straight and Other Juvenile Awareness Programs for Preventing Juvenile Delinquency: A Systematic Review of the Randomized Experimental Evidence. Ann Am Acad Pol Soc Sci. 2003;589:41–62. doi:10.1177/0002716203254693.
115. Thomas J, Harden A, Newman M. Synthesis: Combining results systematically and appropriately. 1st edition. London: Sage; 2012.
116. Fraser MW, Galinsky MJ. Steps in Intervention Research: Designing and Developing Social Programs. Res Soc Work Pract. 2010;20:459–66.
117. Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions. 2008. doi:10.1002/9780470712184.
118. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. Int J Surg. 2010;8:336–41.
119. Tricco AC, Cardoso R, Thomas SM, Motiwala S, Sullivan S, Kealey MR, et al. Barriers and facilitators to uptake of systematic reviews by policy makers and health care managers: a scoping review. Implement Sci. 2016;11:4. doi:10.1186/s13012-016-0370-1.
120. Légaré F, Ratté S, Gravel K, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: Update of a systematic review of health professionals' perceptions. Patient Educ Couns. 2008;73:526–35.
121. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open. 2016;6:e011458. doi:10.1136/bmjopen-2016-011458.
122. Rees R, Oliver K, Woodman J, Thomas J. Children's views about obesity, body size, shape and weight: a systematic review. 2009. http://eppi.ioe.ac.uk/cms/Portals/0/Obesity Views Children R2009Rees.pdf?ver=2010-12-22-121209-040. Accessed 3 Apr 2017.
123. Breen R, Karlson KB, Holm A. Annual Review of Sociology Interpreting and Understanding Logits, Probits, and Other Nonlinear Probability Models. Annu Rev Sociol. 2018;44:4–5. doi:10.1146/annurev-soc-073117.
124. Norton EC, Dowd BE. Log Odds and the Interpretation of Logit Models. Health Serv Res. 2017;:1–20.
125. Muller CJ, Maclehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. Int J Epidemiol. 2014;:962–70. doi:10.1093/ije/dyu029.
126. Vandenbroucke JP, Von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration , for the STROBE Initiative. 2014.
127. Macmillan R. "Distinction" in the third sector. Volunt Sect Rev. 2013;4:39–54. doi:10.1332/204080513X661572.

Chapter 3: What are the barriers and facilitators for third sector organisations (non-profits) to evaluate their services? A systematic review[3]

Anders Malthe Bach-Mortensen, MSc[a]

Paul Montgomery, DPhil, MSc, DipSW[b]

[a] University of Oxford, Department of Social Policy and Intervention; Barnett House, 32 Wellington Square; Oxford; UK

[b] University of Birmingham, Department of Social Policy, Sociology and Criminology, Birmingham, B15 2TT, UK.

Author contributions: The research design of the study was developed by both ABM and PM. ABM conducted the search and screened all the studies, of which PM double-screened a 20% random sample of the studies. All data-extraction, quality appraisal and analysis was conducted by ABM. ABM drafted the manuscript with input from PM. All authors read and approved the final manuscript.

[3] This paper has been published in Systematic Reviews.

Abstract

Background: The third sector is becoming a more common provider of social and health services, but little is known about how third sector organisations (TSOs) evaluate their activities. Past research has reported that the third sector is under increasing pressure to evaluate its impact and performance by government and other commissioning bodies. However, in responding to this increased pressure to undertake evaluation, research suggests that many TSOs struggle to evaluate their activities following the principles of evidence-based practice (EBP). Yet, there has been no systematic effort to investigate *why* the third sector is struggling to provide good quality evidence of its effects.

Methods: This systematic review is reported following PRISMA guidelines. Ten interdisciplinary data-bases were searched using a search string developed following best practice and in consultation with an information systems expert. Included studies were primary research of any research design investigating barriers to and facilitators of the evaluation process of TSOs as identified by practitioners. All studies were quality appraised and the results synthesised as a thematic summary.

Results: Twenty-four studies were included, which mainly investigated TSOs working within health and social services. The thematic summary identified the main barriers for TSOs to undertake evaluation to be related to the (1) lack of financial resources; (2) lack of technical capability and evaluation literacy, and; (3) challenges around identifying relevant evaluation systems and outcome indicators. Key facilitating factors involved (1) getting the appropriate support; (2) having an organisational culture that supports evaluation, and; (3) the motivation to be accountable to stakeholders. These findings were robust to study quality.

Conclusions: This review constitutes the first systematic effort to synthesise existing literature on factors supporting and preventing evaluation by TSOs. The prevalence of factors revolving around the lack of support, resources and clarity on appropriate outcome indicators suggest that many of the identified challenges may be met by applying evidence-based and stakeholder-inclusive strategies to develop shared evaluation requirements. Future efforts should address the application of EBP as part of the commissioning process of TSOs.

Keywords: Evidence-based practice, EBP, barriers and facilitators, evaluation capacity, third sector organisations, TSO, non-profit, evaluation practice, third sector, NGO.

Background

The third sector is under increasing pressure to report on its activities, partly in response to funding/commissioning requirements [1–8]. However, despite the increasing emphasis on impact and performance evaluation by government and commissioners, many third sector organisations (TSOs) fail to evaluate their activities following the principles of evidence-based practice (EBP), which promote rigorous, reproducible and systematic methods [9]. While it is generally acknowledged that monitoring and evaluation according to best practice is important in the delivery of services, there has been little systematic effort to investigate *why* the third sector is struggling to evidence its impact, despite becoming a growing provider of social and health services [10–13].

The 'third sector' is a vague and inconsistently applied term, which covers a range of different types of organisations, such as 'voluntary organisations', 'community-based organisations', 'non-profits' and 'charities', etc., which can each be argued to entail distinct nuances [14, 15].
In the context of this review, we will use the broader term 'third sector organisation' (TSO) and follow the structural/operational definition of TSOs [16], which can be summarised as "…organisations which are formally organised; non-profit distributing; constitutionally independent from the state; self-governing and benefiting from some form of voluntarism (e.g. with volunteer (unpaid) Trustees or Board members or using volunteers in the delivery of services)." [17].

It has been estimated that there are around 166,000 registered TSOs in the UK alone [18], but there is no commonly accepted estimate on the full size of the third sector worldwide. The third sector is becoming an increasingly common provider of public services [7, 19, 20], especially in social care [21, 22] and is considered a central actor in alleviating social issues in most Western countries [7, 23–26]. For example, the primary activities of many TSOs in the UK focus on the

delivery of 'social services' [18]. Further, the UK third sector constitutes a substantial workforce of around 853,000 employees (2016), with an annual spending of £43.3 billion (2014/2015) of which 21% (£9.7 billion) is spent on 'social services' and 10.2% (£4.64 billion) on 'health'. Of its annual income (£45.5 bn 2014/2015), it has been estimated that the government provides £15.3 (33.6%) billion to the third sector, largely through contract-based commissioning. The main beneficiaries of the UK third sector activity are vulnerable groups such as 'children and young people', 'the elderly' and 'people with disabilities' [18].

Recent research and reports demonstrate that many TSOs fail to evidence their activities [27–30]. For example, a 2010 report developed by the Charity Finance Group (CFG) found that only 8% of their sample of 75 TSOs provided external evidence on their impact [28]. More recently, another report which surveyed 1,000 TSOs found that 25% did not evaluate their work at all [27]. While these studies cannot be assumed to be representative of the full population of TSOs, they seem to mirror a growing body of evidence indicating that the sector struggles to adhere meaningfully to the increasing demand for the evaluation of their services [2, 13, 31, 32].

However, few efforts have addressed *why* TSOs that engage in evaluation struggle to demonstrate their impact to stakeholders despite the importance of understanding effectiveness (or possible harms) to vulnerable service-users. To improve the current understanding of third sector practice, research is needed to explore the evaluation process and capacity of TSOs to shed light on what barriers and facilitators they face. To date, there has been no systematic review aggregating the research on the barriers and facilitators for TSOs to undertake evaluation, which constitutes an important research gap since identifying these may be an important step in *improving* current practice and in assessing *how* TSOs may engage in better evaluation practice.

What has been done?

Although there is currently little systematic evidence addressing evaluation practice (or lack thereof) by TSOs, a few studies have investigated the perceived barriers and facilitators of engaging in evaluation [27–29]. This existing research identified factors such as lack of financial resources and support, lack of technical expertise and poor availability of appropriate evaluation and impact tools [27, 28] as being central barriers for TSOs to undertake evaluation. Such barriers are often perceived to be grounded in the failure of funders and regulators to support TSOs in the evaluation process [29, 30, 32]. However, no study has attempted to aggregate the full body of evidence on these factors.

Rather than focusing on the specific challenges of the *evaluation process* of TSOs, past research has focused on the barriers to and facilitators of the *use* and *mobilisation* of research and evidence by TSOs [17] and policy-makers [33, 34]. For example, Hardwick et al. (2015) investigated the knowledge-base in terms of how TSOs utilise research in their activities. This scoping review which considered neither study quality nor importance of the identified factors, found a range of predominantly qualitative studies, which identified various perceived barriers to the use of evidence by TSOs that were mostly related to finance, expertise, resources, and poor collaboration [17]. Perceived facilitators included knowledge mediation, practitioner involvement, and intervention and implementation description. However, the focus of the review was on the knowledge mobilisation of third sector practitioners, and not on the *evaluation process* [17]. While there might be important similarities to the barriers and facilitators associated with the use of evidence and the evaluation process of TSOs, these aspects are fundamentally different, in that TSOs might evaluate their activities without applying evidence as part of one's practice; and vice versa. Further, focusing entirely on the use of evidence, may overlook the issue of evaluation capacity.

Objectives

This systematic review addresses the following question:

- What barriers and facilitators do third sector practitioners identify in relation to evaluating the services their organisations provide?

Methods

We conducted a systematic review in line with the PRISMA guidelines (see appendix A6 for a completed checklist).

Search strategy

High sensitivity was the main aim of the search which was designed to capture all existing research on the barriers and facilitators that TSOs experience in undertaking evaluation. Initially, we conducted preliminary reference list checks and hand-searched/browsed selected journals and data-bases to locate studies that an optimal search would ideally identify. Then, a range of third sector and methodology experts were contacted and invited to provide feedback on the protocol (available upon request) and search strategy (see appendix A5). This then informed the online search in a range of clinical and social science data-bases. The final search string was developed in consultation with an information expert. The included data-bases were:

- ABI/INFORM Global
- Applied Social Sciences Index & Abstracts (ASSIA)
- International Bibliography of the Social Sciences (IBSS)
- MEDLINE®
- PAIS Index
- Policy File Index

- Social Services Abstracts
- Worldwide Political Science Abstracts
- SCOPUS
- Open Grey

Study selection

Included studies were primary research of any research design or systematic reviews that investigated barriers to and facilitators of the evaluation process by TSOs as identified by practitioners. In the context of this review, evaluation is defined as any attempt to quantify or evidence the impact, performance or outcomes of services provided by TSOs [1, 31, 35, 36]. As a result, this review will not be concerned with the distinct nuances between e.g. programme or outcome evaluation and performance measurement etc. [37], effectively assuming that the barriers to different types of evaluation can be generalised. Thus, evaluations such as performance management, evaluation capacity, impact assessment, outcome, programme or service evaluation were considered for inclusion.

Barriers and facilitators are defined as any factors that prevent or support the evaluation capacity of third sector practitioners; they did not need to be the primary outcome of interest of a study to be considered for inclusion.

The third sector and TSOs are denoted inconsistently, and this study considered all synonyms of TSOs, such as 'voluntary organisations', 'community-based organisations', 'non-profits' and 'charity'. Studies that focus on barriers and facilitators in the *use* of evidence by TSOs were excluded, as the focus of this review is on the challenges and needs of TSOs to engage in evaluation. For practical reasons, studies needed to be available in English. There were no date restrictions for studies to be included in the review.

Screening and data-extraction

All abstracts were screened using the online tool Rayyan and data were extracted by ABM with a 20% random sample screened independently by PM. Disagreements were resolved by consensus. The following information was extracted from the included studies:

- Publication year and author
- Study aims
- Study design (time frame/data collection/data analysis)
- Population (type of organisations/services/sample size)
- Results (barriers/facilitators)
- Other results

Data synthesis

The focus of this review is on factors identified by third sector practitioners and included studies were thus either qualitative, survey-based or mixed methods. To synthesise this evidence, the review conducted a thematic summary (sometimes referred to as 'narrative synthesis'), in view of the different types of research studies [38]. This included categorising the identified factors that support and prevent the evaluation capacity of TSOs into themes, in line with previous reviews of this type [33, 39, 40]. This process allowed for a more thorough utilisation of the extracted data, as many of the identified factors could potentially circulate around the same underlying problems, which might be overlooked if the significance of identified factors were only counted by individual frequency. To test the robustness of the thematic framework, and to provide an overview of how the individual studies contributed to the construction of the identified themes, we also constructed a table following Rees et al (2009) [41]. This enabled an assessment of whether certain studies were over- or underrepresented in the thematic framework and helped identify whether e.g. methodology was related to certain themes [38].

Quality assessment

To assess the quality of cross-sectional studies, the review employed the recently developed AXIS checklist [42] (see appendix A1). For the appraisal of interview, ethnography and focus-group studies, an adapted version of the CASP checklist for qualitative studies and the Joanna Briggs Institute QUARI checklist was employed (see appendix A2) [43]. It is important to note that quality appraisal will always entail a degree of value judgement, and that the comparison of the quality of different types of studies (e.g. qualitative and quantitative) may be somewhat arbitrary. Yet, the assessment of study quality arguably allows for a more robust assessment of the findings and of whether the identified factors vary across studies of different quality [38]. Ultimately, appraised studies were organised into three categories – high, medium and low. The basis of the quality ranking was informed by the number of checklist items fulfilled by the individual studies, but the final assessment of quality was based on an overall judgement of the value of the individual findings, considering the methodology of the included studies.

Results

Prior to the formal search, 10 potentially eligible studies had been identified through preliminary scoping of reference lists and hand-searches. The final search located 3406 studies of which 2729 unique studies were identified after removing duplicates. Of these, 2660 were excluded after screening the abstracts. Sixty-nine studies were then read in full-text of which 45 were excluded because: not focusing on the evaluation process of TSOs (n=18), not representing the experiences of TSO practitioners (n=10), not being primary research (n=9), focusing on the collaboration between researchers and community programmes (n=5), or focusing on the implementation of evidence-based procedures rather than evaluation of services (n=3). Twenty-four studies were included in the final synthesis.

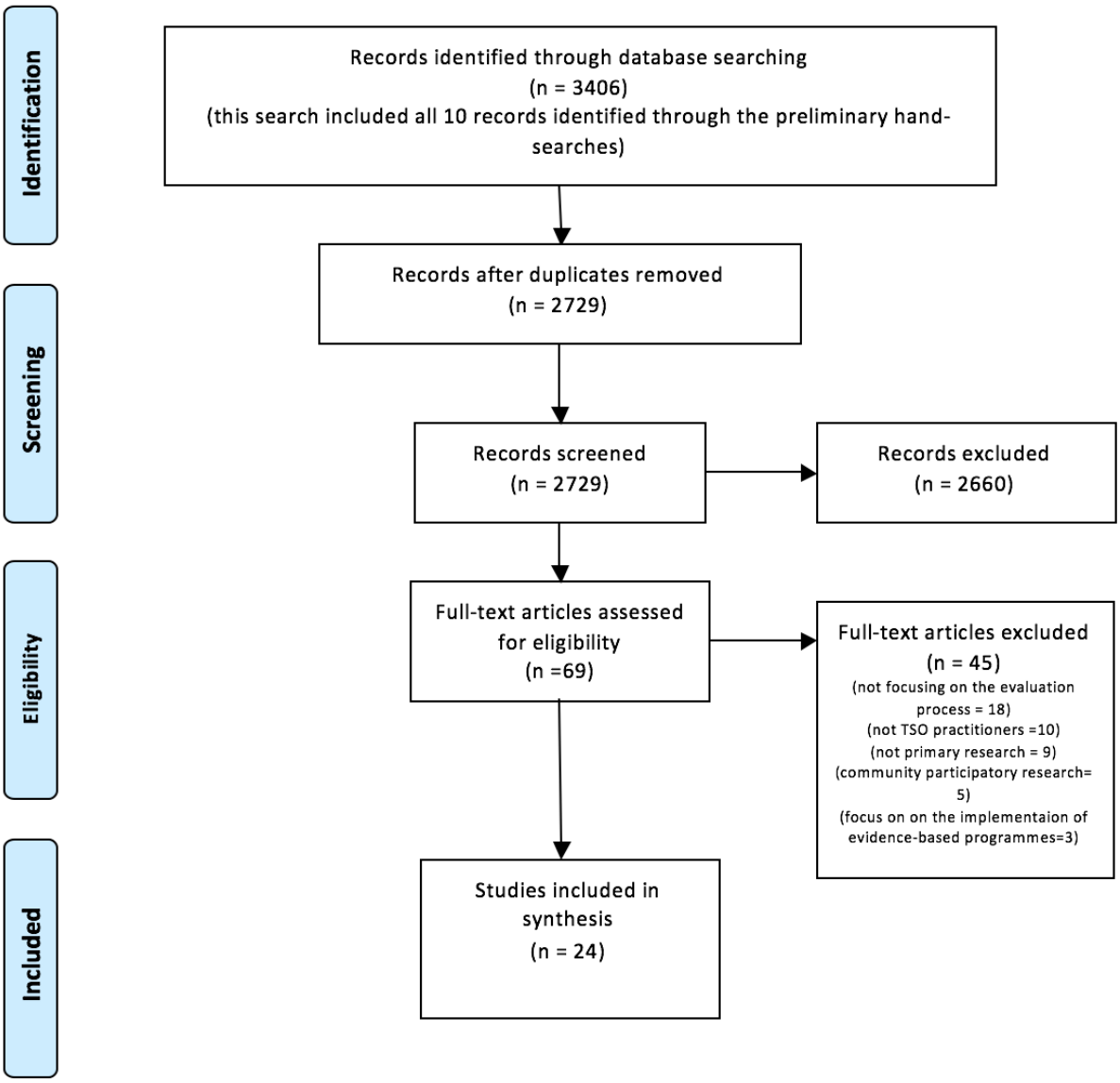


Figure 3: PRISMA diagram.

Characteristics of included studies

The full characteristics and extracted data of the included studies can be found in the online additional file 3. Of the twenty-four included studies, nine were reports written by and/or for third sector organisations, and fifteen were articles published in peer-reviewed journals. Nine studies (37.5%) were purely quantitative, nine purely qualitative (37.5%) and six employed mixed methods (25%). The clear majority of studies (70.8%) were conducted in North America, followed by the UK (20.8%). Only one study (4.2%) was conducted entirely in a low and middle-income country (LMIC), and one study (4.2%) was conducted in multiple contexts - US,

Columbia, Guam and Puerto Rico. Eight of the studies (33.33%) were published in the period between 2000-2009, and sixteen studies (66.7%) after 2010.

Samples of TSOs in included studies

Thirteen of the included studies (54.2%) had a sample size above 100 TSOs and only five (20.8%) included less than 20 TSOs. Of the studies with a sample < 100, three had a sample of 1-5 TSOs, two contained 6-20 [44, 45], four a sample of 21-50 [46–49], and two a sample of 51-99 [50, 51]. Of the studies with a sample > 100 TSOs, seven included 100-499 TSOs [28, 52–57], two included 500-999 TSOs [29, 58], and four studies a sample of a 1000 or more TSOs [27, 59–61].

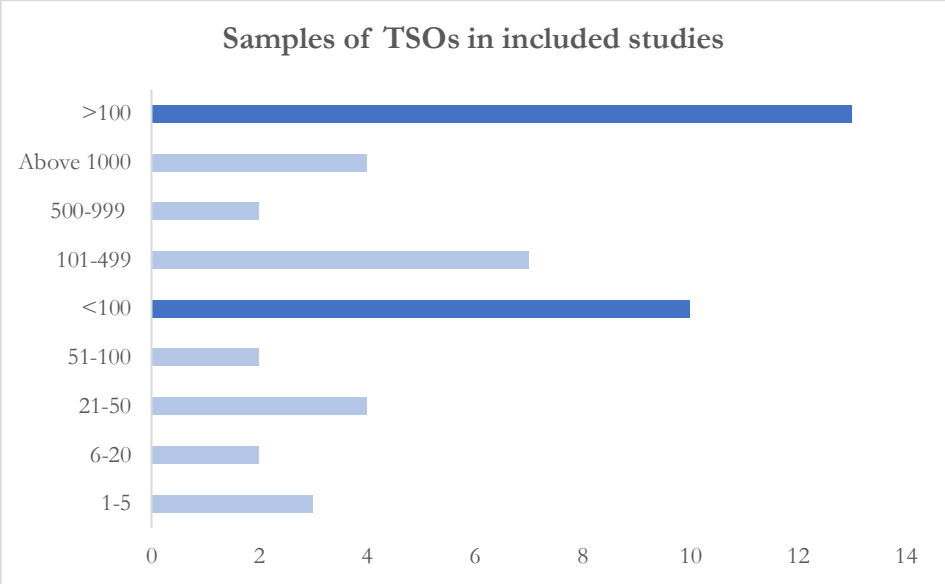


Figure 4: Samples of TSOs in included studies.

Size of TSOs

Seventeen of the studies (70.8%) included a mixed population TSOs in terms of income [27–29, 44, 45, 47, 48, 50, 52–55, 57–61], of which eight had a self-reported bias towards larger organisations [27, 28, 45, 48, 53, 58, 60, 61]. Two studies [62, 63] only included small TSOs (annual income below £100,000), one study [49] focused on small- to medium-sized TSOs

(£100,000-£1million), and one report [64] looked at medium to large organisations (above £1million). Three studies [46, 51, 56] did not specify the income of their included organisations, as their focus was on the evaluation capacity of organisations in the context of the implementation of specific programmes (e.g. 'Healthy Tomorrow' [56]).

Purpose of TSOs

Most of the included studies (70.8%) investigated TSOs working primarily in social and human services (e.g. health, children and families, HIV/aids), followed by six studies (25%) that did not specify the included population of TSOs, and one study (4.2%) working with environmental preservation organisations. Seven of them (29.2%) included mixed types of organisations in their analysis, of which organisations delivering social services constituted the biggest proportion. The biggest reported sub-categories of social services were organisations working with children and families (25%), HIV/aids service organisations (20.8%) and education (16.7%).

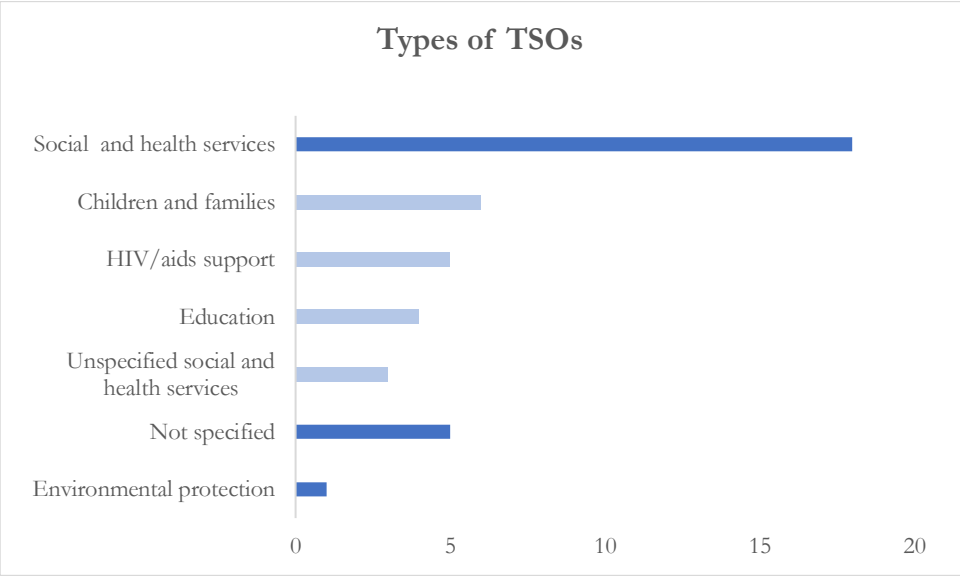


Figure 5: Type of service delivery of the organisations included in the studies.

Type of evaluation

All studies investigated evaluation in some way, but used different terminology on the specific *type* of evaluation. Six studies (25%) investigated 'performance evaluation', five studies (20.8%)

looked at 'evaluation capacity', eight studies (33.3%) focused on 'evaluation practice', four studies (16.7%) on 'programme evaluation', and one study (4.2%) focused on 'impact evaluation'.

Quality of included studies

All 24 studies were subject to quality appraisal. Three studies (12.5%) were deemed to be of 'high quality' [32, 51, 64], fifteen (62.5%) of 'medium quality' [27, 29, 44–48, 50, 52, 53, 56, 57, 59–61], and the remaining six studies (25%) [28, 49, 55, 58, 62, 63] were assessed to be of 'low quality'. Of the studies that employed mixed methods, the appraisal focused on the methodology which identified the factors operating as barriers and facilitators. Most of the mixed methods studies were predominantly surveys [28, 29, 32, 55, 57, 59], but included interviews and/or focus-groups to inform or validate the findings, and were consequently appraised using cross-sectional criteria.

Identified barriers and facilitators

All 24 studies investigated factors operating as barriers for TSOs to evaluate their activities, whereas only 15 of those studies included factors that facilitate evaluation among TSOs. Only two studies [45, 51] identified the factors operating as barriers and facilitators as being reciprocal, whereas all other studies found them to be distinct. The synthesis identified 30 unique factors operating as barriers and 26 factors operating as facilitators. Interestingly, many of the identified barriers and facilitators did not mirror each other, as is often the case in reviews of this type [33, 40].

The top 3 cited factors operating as barriers to engagement in evaluation focused on the lack of expertise and internal capacity (17/24), mismatch between funder requirements and what TSOs perceived to be appropriate evaluation goals (16/24), and the lack of financial resources to conduct evaluation (14/24). The most reported factors operating as facilitators to evaluation included funders requiring evaluation (6/15), involving stakeholders in identifying relevant

outcome indicators and evaluation goals (5/15), and having appropriately trained staff to undertake evaluation (5/15).

Table 1: Top 3 cited barriers and facilitators.

Top 3 reported barriers	Top 3 reported facilitators
Lack of expertise and internal capability (17/24)	Funder requirements (6/15)
Mismatch between funders requirements and appropriate goals (16/24)	Involvement of stakeholders to identify outcome indicators and evaluation goals (5/15)
Lack of financial resources (14/24)	Training of staff and evaluation literacy (5/15)

Weight of factors

While most of the studies did not rank the identified factors hierarchically, some studies did argue that certain factors were more significant than others, which was based on that factor being more frequently reported by respondents. Of the studies that did rank identified barriers, *lack of time* was the factor that most studies argued to be the *main* barrier to engage in evaluation (n=6) [29, 31, 54, 58, 60, 61]. For facilitators, the factor most often reported as the *main* facilitator to undertake evaluation was having an organisational management that supports and requires evaluation (n=2) [59, 60].

Themes

Many of the reported barriers and facilitators were described differently, but revolved around the same underlying problem (e.g. the lack of financial and staff resources may reflect similar constructs). To approach the fact that many factors revolve around similar fundamental issues, we organised all the identified factors into overarching categories (Table 2), in line with the purpose of this type of review [33, 39, 40].

As Table 2 shows, the two biggest categories preventing TSOs from undertaking evaluation were factors associated with the lack of resources (e.g. lack of money, time and staff) (n=36) and the lack of technical capability and skills to evaluate (n=36). The third biggest theme was challenges

related to utilising evaluation systems and identifying relevant impact and outcome indicators (n=27). Factors linked to organisational culture was also a significant theme (n=26), which e.g. included staff resistance to evaluation and lack of support to engage in evaluation from the board and leadership. Another significant theme included factors associated with funder requirements (n=25), which was mainly constituted by the perceived inappropriateness of the evaluation and reporting requirements set by commissioners.

Table 2: Factors operating as barriers and facilitators across identified themes.

Factors operating as barriers	Times cited
Factors related to lack of resources	**36**
Financial resources	14
Staff resources	6
Lack of time	6
Lack of resources to hire external evaluators	3
Resources (unspecified)	6
Staff turnover	1
Factors related to technical capability and evaluation skills	**36**
Lack of expertise and internal capability	17
Problems with data-collection and analysis	11
Inability to utilise existing data	3
Difficulty conceptualising and designing evaluation	1
Technical challenges	4
Challenges in utilising evaluation systems and identifying outcome indicators	**27**
Difficulty developing and using evaluation tools	5
Lack of integrated systems to collect and analyse data	9
Challenges in identifying accepted outcome and impact indicators	13
Factors related to organisational culture and management	**26**
No perceived benefit to conduct evaluation	1
Staff resistance to evaluation	10
Perceived compromise between evaluation and service delivery	3
Lack of evaluation strategies and planning	3
Lack of feedback between board and management staff	1
Low prioritisation of evaluation	3
Lack of support from board and leadership	4
Evaluation not part of everyday practice	1
Factors related to funder requirements	**25**
Lack of funders requirements and support	3
Differing requirements from different funders	1
Mismatch between funders requirements and appropriate goals	16
Poor proportioning of size of charities and funders requirements	1
Micro-management by donors	1
Funding insecurity (funding circles incentivising focusing on immediate outputs rather than long terms outcomes)	3
Other	**5**

Confidentiality of data	3
Lack of cooperation with stakeholders	2

Factors operating as facilitators	
Factors related to receiving support to evaluate	**19**
Partnering with evaluation experts	3
Partnering with organisations working with similar activities	1
Technology availability and literacy to collect and analyse data	4
Benchmark data availability	2
Training of staff and evaluation literacy	5
Workable evaluation tools	2
Having sufficient resources to evaluate	2
Factors related to organisational culture and management	**18**
Understanding internal processes	3
To embed evaluation as part of everyday practice	1
Improve allocation of resources	1
Support from board and leadership	4
Have in-house evaluation staff	3
Having clear goals and evaluation strategies	2
Staff support	3
Positive perception of evaluation	1
Factors related to the motivation to be accountable	**17**
Involvement of stakeholders to identify outcome indicators and evaluation goals	5
The motivation to influence policy	1
Compare work and outcomes to others doing similar work	1
The motivation to inform the sector as whole	1
Improve targeting of beneficiaries	1
Identify new approaches	1
Ensuring control and legitimacy of activities to stakeholders	3
The motivation to demonstrate and improve effectiveness of services	4
Factors around funder requirements and regulations	**7**
Funder requirements	6
Regulation requirements	1
Factors around economic sustainability	**5**
Using evaluation to be eligible for funding opportunities	3
Using evaluation as marketing	2

The most reported category of facilitating factors focused on getting the appropriate support to undertake evaluation (n=19), which included e.g. receiving expert support in evaluating services and receiving staff training to ensure the appropriate literacy to run evaluations. The second most reported category were factors related to organisational culture and management, which e.g. included staff and management support to undertake evaluation. The third most salient category of facilitating factors was motivation to be accountable (n=17), which mainly revolved around including stakeholders in defining outcome and impact criteria, as well as ensuring

control and effectiveness of services. Lesser reported factors facilitating evaluation were those related to the requirements set by funders and regulators (n=7) and economic sustainability (n=5).

Table 3: Overview of the contribution of individual studies on the identified themes of barriers.

Study	**Study methodology**	**Sample size (TSOs)**	Factors related to lack of resources	Factors related to technical capability and evaluation skills	Challenges in utilising evaluation systems and identifying outcome indicators	Factors related to organisational culture and management	Factors related to funder requirements
Hoefer (2000)	Survey	91	X	X			X
Mitchell and Berlan (2016)	Survey	311	X	X	X		X
Kang et al. (2012)	Survey	38	X	X	X	X	
Ogain et al. (2012)	Survey	1,000	X	X	X	X	X
Williams-Gray (2016)	Survey	265	X	X	X	X	
Morariu et al. (2012)	Survey	535	X	X		X	X
Ruch-Ross et al. 2008	Survey	126	X	X	X		
Reed and Morariu (2010)	Survey	1,072	X	X		X	
Morariu et al. (2016)	Survey	1,125	X	X			X
Big Lottery Fund (2013)	Mixed methods (surveys, interviews and focus-groups	404	X	X	X	X	X
Andrei et al. (2012)	Mixed methods (surveys and focus groups)	398	X	X		X	
Carman and Fredericks (2010)	Mixed methods (surveys and follow-up interviews)	189	X	X	X	X	X
Breckell et al. (2010)	Mixed methods (focus groups and surveys)	164	X		X		X
Ellis (2009)	Mixed methods (survey, interviews and workshops)	682	X		X	X	X

Hall et al. (2003)	Mixed methods (surveys and focus groups)	1,965	X	X	X		X
Bourgeios et al. (2015)	Interviews	4	X	X	X		
Minich et al. (2006)	Interviews	24	X	X	X		
Kegeles et al. (2005)	Interviews	44	X	X	X	X	
Newcomer et al. (2013)	Interviews	12		X	X		X
Carnochan et al. (2014)	Interviews and focus-groups	4 (36 interviewees)		X	X	X	
Moxham (2007)	Interviews	4	X		X		X
Thomson (2011)	Interviews	18	X	X	X		
Atkintobi et al. (2012)	Interviews	24	X		X		
Napp et al. (2002)	Interviews	61	X	X	X	X	X

The table reflects how each study contributed to the construction of the final themes of factors operating as barriers. However, it should be noted that each study can contribute with multiple factors to the same theme. Shading indicates studies that were rated to be of low quality.

Table 4: Overview of the contribution of individual studies on the identified themes of facilitators.

Study	**Study methodology**	**Sample size (TSOs)**	Factors related to receiving support to evaluate	Factors related to organisational culture and management	Factors related to the motivation to be accountable	Factors related to funder requirements and regulation	Factors around economic sustainability
Hoefer (2000)	Survey	91			X	X	X
Mitchell and Berlan (2016)	Survey	311	X	X	X	X	X
Kang et al. (2012)	Survey	38	X	X			
Ogain et al. (2012)	Survey	1,000	X	X	X		
Morariu et al. (2016)	Survey	265	X	X	X		
Breckell et al. (2010)	Mixed methods (focus groups and surveys)	164		X			
Ellis (2009)	Mixed methods (survey, interviews and workshops)	682		X			

Hall et al. (2003)	Mixed methods (surveys and focus groups)	1,965		X		X	
Big Lottery Fund (2013)	Mixed methods (surveys, interviews and focus-groups)	404	X	X	X	X	
Minich et al. (2006)	Interviews	24	X	X			
Kegeles et al. (2005)	Interviews	44		X	X	X	
Carnochan et al. (2014)	Interviews and focus-groups	4 (36 interviewees)		X	X		
Thomson (2011)	Interviews	18	X	X			
Atkintobi et al. (2012)	Interviews	24	X	X			
Napp et al. (2002)	Interviews	61	X	X		X	

The table reflects how each study contributed to the construction of the final themes of factors operating as facilitators. However, it should be noted that each study can contribute with multiple factors to the same theme. Shading indicates studies that were rated to be of low quality.

Robustness of themes

Table 3 and 4 illustrate that the salience of themes corresponded to how representative the themes were of the included studies. However, 'organisational culture and management' was an exception, in that it represented more of the included studies, compared to other top categories of facilitators. Notably, it was only survey studies which did not recognise 'challenges around identifying appropriate outcome indicators' as a barrier [50, 55, 58, 60, 61], and only two studies, both qualitative, did not identify 'resources' as a barrier to evaluation [44, 64]. However, the tables do not reflect any clear pattern to suggest that methodology or sample size systematically affect how the included studies contribute to the specific themes.

Sensitivity analysis

To test for the sensitivity of the results according to study quality, a sensitivity analysis was conducted by including only the 'high' or 'medium' quality studies. The top three main reported factors operating as facilitators and barriers remained the same (only two of the studies of low

quality included facilitating factors) (see appendix A3). Looking at how excluding the studies of low quality effect the thematic categories (see appendix A4), the most reported themes still revolved around technical capability and evaluation skills (n=28), factors associated with resources (n=26) and factors associated with evaluation systems and outcome indicators (n=21). The main categories of facilitating factors similarly remained the same after removal of low quality studies suggesting that the main identified categories are robust to this analysis.

Discussion of findings

Summary of findings

This systematic review constitutes the first systematic attempt to synthesise existing research on what factors support or prevent evaluation for TSOs. The search identified fifteen peer-reviewed publications and nine reports that can be classified as grey literature. The review identified the main factors preventing TSOs to undertake evaluation to revolve around the lack of financial resources, technical capability and evaluation literacy, and challenges around identifying relevant evaluation systems and outcome indicators. The main facilitators involved getting appropriate support to evaluate, having an organisational culture and management in favour of evaluation, and the pursuit of accountability. These findings were robust across different levels of study quality and were largely representative of the full body of included studies.

Implications of findings

The thematic synthesis demonstrated that the main barriers for TSOs to undertake evaluation was related to organisational *capacity* and *capability*. One may think of capacity as 'how much' an organisation can do, relating to issues around resources, and capability as 'how well' an organisation can operate, thus relating to the skills, knowledge and confidence of organisations [65, 66]. These issues were reflected in the most reported categories of factors operating as

barriers; 'factors related to lack of resources' (capacity), and 'factors related to technical capability and evaluation skills' (capability).

Importantly, these findings suggest that many TSOs are not receiving the appropriate capacity *or* capability support in evaluating their services. The significance of these barriers is further illustrated by the most reported categories of facilitating factors being to receive the appropriate support to undertake evaluation. The prevalence of these categories implies a rather unfortunate situation in which TSOs are faced with a growing pressure to evidence their performance to secure grants or contracts, but without being subject to the necessary support to undertake meaningful evaluation.

This situation might shed light on the increasingly reported tendency that many TSOs customise their evaluation procedures not to improve their service delivery, but rather to satisfy funding bodies and to meet the increased pressure of performance measurement [27, 29, 32, 67, 68]. Having an incentive structure which encourages TSOs to shape evaluation entirely according to funders requirements is greatly problematic, as it may override what is arguably the most important pursuit of evaluation, i.e. to *monitor* and *improve* practice [69, 70]. Further, if those evaluation criteria are flawed or inappropriate to the activities to which they apply, the commissioning process may facilitate a market that incentivises poor and quick evaluations.

However, it was not only *external* barriers and facilitators that were identified as being central for TSOs to undertake evaluation. The themes of organisational culture and management were consistent across factors operating as both barriers and facilitators; particularly regarding staff resistance/support and organisational motivation (or lack thereof) to engage in evaluation. This finding is, for example, mirrored in a recent report by the Big Lottery Fund, which found that despite allowing its grantees to spend up to 10% of the allocated funding on evaluation, half of

the surveyed TSOs (n=404) only allocated between 1% and 5% of their budget to evaluation activities [57]. Improvement of evaluation capacity and capability among TSOs thus not only requires increased external support in terms of funders allowing for higher evaluation spending. This point is further exhibited by one of the most cited categories of facilitating factors being 'organisational culture and management' (n=18) and 'the motivation to be accountable' (n=17), suggesting that it is essential for TSOs to demonstrate *internal* motivation and will to mobilise existing resources towards improving evaluation capacity and capability [47, 51, 64].

This review also demonstrated that TSOs face considerable challenges regarding the unclear terminology and fluctuating understanding of terms such as 'evaluation' and 'performance management'. This challenge is reflected by the second most reported barrier being the perceived mismatch between the requirements set by funders and what practitioners deemed to be valuable pursuits (16/24), and by one of the most reported facilitators being to involve stakeholders in the process of identifying outcome and impact indicators (5/15). This implies that when it comes to choosing *how to evaluate* and *what to measure*, there is poor consensus among funders, practitioners and other stakeholders about what constitutes 'good practice'. Such perceived lack of consensus points to an absence of stakeholder collaboration and inclusion in defining and planning the evaluation requirement of third sector services.

Ways forward

The most cited groups of factors facilitating TSOs to undertake evaluation was getting the appropriate support to evaluate (n=19), in which 'financial support' was only identified twice as a facilitating factor under this theme. The low reporting of financial resources as a facilitating factor is noteworthy, considering that 'lack of resources' (n=36) was among the most cited categories of factors operating as barrier, along with the lack of evaluation skills and capability (n=36). This suggest that increasing financial support for TSOs to evaluate might be a necessary

factor to improve evaluation practice, but that it is not sufficient in isolation. Rather, the analysis demonstrates that TSOs need improved support in the entire *evaluation process*, which would require attention to both capacity (more resources) and capability (e.g. staff training and collaboration with experts and researchers).

Further, there seems to be a need for clearer and more commonly accepted outcome indicators and evaluation guidelines that can be effectively utilised by third sector practitioners. Currently, there is a wide range of available evaluation frameworks and impact assessment tools made for and by TSOs [4, 71–76]. These tools include social returns on investment (SROI), social auditing, and other types of outcome measurement [1, 73, 74, 77]. However, there is limited reliable data on the *use* and *uptake* of the various existing impact evaluation frameworks, and there are not any universal guidelines and consensus as to how different TSOs should apply them [1, 4, 26, 27, 29, 73, 77]. The findings of this review suggest that the failure to provide clear and consensus-based guidelines on the *appropriateness* of different types of evaluation procedures operate as a central barrier for TSOs to undertake evaluation.

The role of consensus and stakeholder inclusion is an essential aspect in adhering to the principles of EBP [78–80]. For example, when deciding on the appropriateness of new reporting and evaluation guidelines in clinical practice, methods such as the Delphi technique are widely considered to be best practice, and served as the main tool to develop the CONSORT guidelines and its extensions [81–83]. While third sector scholars have continuously argued against the notion of a universal performance measurement, this review supports the idea that more efforts are needed to determine consensus-based and manageable criteria for outcome indicators and effectiveness in the third sector [8, 84]. Importantly, this does not necessarily require the adherence to strict or 'one-size-fits-all' outcome or performance criteria. Rather, one might utilise consensus-based techniques such as the Delphi method to develop stakeholder inclusive

processes, from which an evaluation strategy can be defined by both service-users, practitioners and funders. Further, engaging with stakeholders would constitute a central step in deciding what types of organisational support are necessary to ensure that TSOs can adhere to appropriate evaluation criteria. However, there is currently little research that addresses how consensus procedures from EBP can be utilised in the third sector.

Limitations of this study

The main purpose of this study was to aggregate all available research addressing the current evaluation practice of TSOs, and thus serves as an inclusive exploration of existing research investigating factors affecting third sector evaluation capacity and capability. Like most reviews of this type, the study depends on a form of 'vote-counting', in that the findings are constituted by the aggregation of factors reported as barriers and facilitators. However, by organising the factors into identified themes, the reported factors should be less sensitive to the, arguably, arbitrary difference between evaluation expertise and experiencing challenges with collecting and analysing data. Further, the three most reported barriers as well as the ranking of the themes remained consistent when removing the 'low quality' studies, which points to the overall robustness of the analysis.

It is important to note that many of the surveys suffered from either very low response-rates [31, 60, 61], non-reporting of the response-rates [28, 29, 55, 58], and the tendency to have a bias towards including larger organisations [27, 28, 45, 48, 53, 58, 60, 61]. Also, in the cross-sectional studies, TSOs were restricted to the factors that were included (or not included) in the survey questionnaires. This constitutes a considerable limitation, as those factors might not have been the same if following an inductive or bottom-up qualitative approach. However, many of the cross-sectional studies included focus-groups and interviews to ensure that the questions were reflective of issues that were recognised by TSO practitioners.

A more substantial limitation of this review is that different types of organisations might be subject to different factors preventing or supporting their evaluation capacity, which compromises the generalisability of the findings. For example, one of the included studies found a separate set of barriers between NGOs operating in Egypt and in Columbia [44]. Also, several studies found that smaller TSOs experience greater challenges in evaluating their services [27, 47, 54, 55]. To explore this tendency, one study [54] ran a logistic regression of different organisational characteristics to look for correlations in what factors were reported. This analysis found six organisational characteristics to be significantly associated with the inclination to report different types of barriers. However, when exploring these characteristics through cluster analysis, the study found that the size of organisations, type of services and the funding sources did not explain any variability between the reported barriers of the different clusters [54].

The included population of this review was predominantly TSOs from Western countries delivering social and health services. However, given that third sector spending tends to be higher in Western TSOs, these findings may be generalisable to a larger population of organisations [85, 86]. Moreover, the delivery of social services often include the potential to cause harm to the service-user [11, 12, 87, 88], and recent research suggests that TSOs delivering social services are more likely to be publically funded [89], thus making the evaluation capacity of this type of TSOs particularly important to investigate.

Future research

This review focused on the perspectives of third sector practitioners, following the assumption that they would provide the most reliable accounts regarding factors influencing the evaluation capacity of TSOs. Future research might expand this focus to a wider group of stakeholders, which might shed light on how commissioning decisions are made in practice [90, 91]. Also, one

might argue that the focus on factors operating as barriers and facilitators is restrictive, and future research might benefit from taking a broader approach by focusing on the broad *experiences* of TSOs to undertake evaluation. Last, this review did not investigate factors affecting the capacity for TSOs to *implement* evidence-based programmes and interventions, which is often acknowledged to be a central limitation in the third sector delivery of social and clinical work [13, 92]. Building on this review, future efforts should be made to investigate the barriers and facilitators related to the *implementation* of evidence-based interventions and programmes by TSOs.

Conclusions

The main implication identified in this systematic review is the apparent willingness of funders and regulators to enforce evaluation on TSOs without offering support and consistent guidelines as to how evaluation should be undertaken. To address this problem, it is central not to focus on individual barriers (such as financial resources), but to consider what support is necessary to ensure that TSOs have the appropriate capacity and capability to undertake evaluation [65].

This may be achieved by having a procedure for consensus, to determine, first of all, what evaluation criteria different types of TSOs require, and, equally important, what types of supports will enable TSOs to conduct meaningful evaluation. Without the assurance that evaluation requirements are based on 'best practice' and stakeholder perspectives, the utility of evaluation (i.e. to improve practice and prevent iatrogenic effects) is jeopardised, and the service-users of the third sector are put at risk. To improve the current evaluation practice of the third sector, guideline development procedures employed by EBP may enable improved and consensus-based guidance that incorporates stakeholder perspectives with the best available research [83].

Most research on the performance of TSOs concludes that there is insufficient knowledge about current activities to make general inferences about the effectiveness of the third sector delivery of services [7, 17, 21]. This means that is difficult to assess with confidence whether the increased third sector delivery of public and social services is indeed effective and worthwhile [7]. This paper is the first systematic review to investigate what factors obstruct and promote evaluation by TSOs, and as such represents an important step in trying to *improve* current practice. The findings of this review may inform future efforts to investigate how current funding requirements can be adapted to facilitate the necessary evaluation capacity and capability for TSOs to evidence activities according to best practice.

List of abbreviations:

TSO: Third sector organisation, EBP: Evidence-based practice.

Acknowledgements: The authors would like to thank librarian John Southall at the Social Science Library, University of Oxford, for providing helpful feedback to the search strategy employed by this study. We are also indebted to Professor Jane Barlow and Mads Roke Clausen for providing helpful comments and improvements to an earlier draft of the manuscript.

References

1. Barman E. What is the bottom line for nonprofit organizations? A history of measurement in the british voluntary sector. Voluntas. 2007;18:101–15. doi:10.1007/s11266-007-9039-3.
2. Carman JG. Evaluation practice among community-based organizations: Research into the reality. Am J Eval. 2007;28:60–75. doi:10.1177/1098214006296245.
3. Bovaird T, Dickinson H, Allen K. Commissioning across government: review of evidence. TSRC Research Report 86 (Birmingham, University of Birmingham). 2012. https://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/reports/research-report-86-commissioning-across-govt.pdf. Accessed 25 Oct 2016.
4. Harlock J. Impact measurement practice in the UK third sector: A review of emerging evidence. TSRC Working Paper 106 (Birmingham, University of Birmingham). 2013. http://epapers.bham.ac.uk/1800/1/WP106_Impact_measurement_practice_in_the_UK_third_sector_-_Harlock%2C_July_2013.pdf. Accessed 7 Apr 2016.
5. Harlock J, Metcalf L. Measuring impact: prospects and challenges for third sector

organisations. Volunt Sect Rev. 2016;7:101–8. doi:10.1332/204080516X14534734765005.
6. Kendall J, Knapp M. Measuring the Performance of Voluntary Organizations. Public Manag Rev. 2000;2:105–32. doi:10.1080/14719030000000006.
7. Macmillan R. The third sector delivering public services: an evidence review. TSRC Working Paper 20 (Birmingham, University of Birmingham). 2010. https://www.birmingham.ac.uk/generic/tsrc/research/service-delivery/wp-20-public-services-evidence-review.aspx. Accessed 7 Apr 2016.
8. Moxham C. Understanding third sector performance measurement system design: a literature review. Int J Product Perform Manag. 2014;63:704–26. doi:10.1108/IJPPM-08-2013-0143.
9. Sackett D. Evidence-based medicine : how to practice and teach EBM. Churchill Livingstone; 2000. doi:10.1016/S1031-170X(97)80036-0.
10. Chalmers I. If evidence-informed policy works in practice, does it matter if it doesn't work in theory? Evid Policy. 2005;1:227–42.
11. Chalmers I. Trying to do more Good than Harm in Policy and Practice: The Role of Rigorous, Transparent, Up-to-Date Evaluations. Ann Am Acad Pol Soc Sci. 2003;589:22–40. doi:10.1177/0002716203254762.
12. McCord J. Cures That Harm: Unanticipated Outcomes of Crime Prevention Programs. Ann Am Acad Pol Soc Sci. 2003;587:16–30.
13. Despard MR. Challenges in Implementing Evidence-Based Practices and Programs in Nonprofit Human Service Organizations. J Evidence-Informed Soc Work. 2016;13:505–22. doi:10.1080/23761407.2015.1086719.
14. Alcock P. A strategic unity: defining the third sector in the UK. Volunt Sect Rev. 2010;1:5–24.
15. Evers A, Laville J-L. Defining the third sector in Europe. Edward Elgar Publishing; 2004.
16. Salamon LM, Anheier K. In search of the non-profit sector. I: The question of definitions. 1992. http://download.springer.com/static/pdf/414/art%253A10.1007%252FBF01397770.pdf?originUrl=http%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252FBF01397770&token2=exp=1491909555~acl=%252Fstatic%252Fpdf%252F414%252Fart%2525253A10.1007%2525252FBF01. Accessed 11 Apr 2017.
17. Hardwick R, Anderson R, Cooper C. How do third sector organisations use research and other knowledge? A systematic scoping review. Implement Sci. 2015;10:84. doi:10.1186/s13012-015-0265-6.
18. NCVO. NCVO UK Civil Society Almanac. 2017. https://data.ncvo.org.uk/. Accessed 5 Jan 2017.
19. Moxham C, Boaden R. The impact of performance measurement in the voluntary sector - identification of contextual and processual factors. Int J Oper Prod Manag. 2007;27:826–45. doi:10.1108/01443570710763796.
20. Office for National Statistics. Measuring outcomes for public service users. 2010.
21. Dickinson H, Allen K, Alcock P, Macmillan R, Glasby J. The Role of the Third Sector in Delivering Social Care. NIHR School for Social Care Research. 2012. doi:978-0-85328-443-7.
22. Barron DN, West E. The quasi-market for adult residential care in the UK: Do for-profit, not-for-profit or public sector residential care and nursing homes provide better quality care? Soc Sci Med. 2017;179:137–46. doi:10.1016/j.socscimed.2017.02.037.
23. Hogg E, Baines S. Changing Responsibilities and Roles of the Voluntary and Community

Sector in the Welfare Mix: A Review. Soc Policy Soc. 2011;10:341–52. doi:10.1017/S1474746411000078.
24. Kelly J. Reforming public services in the UK: Bringing in the third sector. Public Adm. 2007;85:1003–22. doi:10.1111/j.1467-9299.2007.00680.x.
25. Cabinet Office. Open Public Services White Paper. 2011. http://www.cabinetoffice.gov.uk/sites/default/files/resources/open-public-services-white-paper.pdf%5Cnhttp://www.openpublicservices.cabinetoffice.gov.uk/.
26. Anheier HK. Nonprofit Organizations: Theory, Management, Policy. Routledge; 2005. http://books.google.com/books?id=s_uzg-KoVyIC&pgis=1.
27. Ógáin EN, Lumley T, Pritchard D. Making an Impact: Impact measurement among charities and social enterprises. 2012; October:1–73. http://www.thinknpc.org/publications/making-an-impact/. Accessed 10 Aug 2016.
28. Breckell P, Harrison K, Robert N. Impact Reporting in the UK Charity Sector. 2010. http://www.gaugeni.co.uk/sites/default/files/Policy_cfdg_impact_reporting_UK_charity_sector.pdf%5Cnhttp://www.gaugeni.co.uk/sites/default/files/Policy_cfdg_impact_reporting_UK_charity%5Cn_sector.pdf.
29. Ellis J, Gregory T. Accountability and learning: developing monitoring and evaluation in the third sector. CES Research Report. 2008. http://www.juconicomparte.org/recursos/cesresearchpaper-accountabilityandlearning-540-548_MZZ1.pdf. Accessed 1 Mar 2017.
30. Carman JG. Evaluation practice among community-based organizations: Research into the reality. Am J Eval. 2007;28:60–75. doi:10.1177/1098214006296245.
31. Mitchell GE, Berlan D. Evaluation and Evaluative Rigor in the Nonprofit Sector. Nonprofit Manag Leadersh. 2016;27:237–50.
32. Carman JG, Fredericks KA. Nonprofits and Evaluation: Empirical Evidence From the Field. In: Nonprofits and evaluation. New Directions for Evaluation. 2008. p. 51–71.
33. Oliver K, Innvar S, Lorenc T, Woodman J, Thomas J. A systematic review of barriers to and facilitators of the use of evidence by policymakers.(Research article). BMC Health Serv Res. 2014;14:2.
34. Innvar S, Vist G, Trommald M, Oxman A. Health policy-makers' perceptions of their use of evidence: a systematic review. J Health Serv Res Policy. 2002;7:239–44.
35. Wikipedia F. Performance measurement. Evaluation. 2009;29:1–8. doi:10.1007/BF03249635.
36. Moxham C. Performance measurement Examining the applicability of the existing body of knowledge to nonprofit organisations. Int J Oper Prod Manag. 2009;29:740–63.
37. Hatry HP. Sorting the Relationships Among Performance Measurement, Program Evaluation, and Performance Management. New Dir Eval. 2013;2013:19–32.
38. Thomas J, Gough D, Oliver S. Synthesis: Combining results systematically and appropriately. An Introd to Syst Rev. 2012;:179–226.
39. Gravel K, Légaré F, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: a systematic review of health professionals' perceptions. Implement Sci. 2006;1:16. doi:10.1186/1748-5908-1-16.
40. Légaré F, Ratté S, Gravel K, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: Update of a systematic review of health professionals' perceptions. Patient Educ Couns. 2008;73:526–35.
41. Rees R, Oliver K, Woodman J, Thomas J. Children's views about obesity, body size, shape

and weight: a systematic review. 2009. http://eppi.ioe.ac.uk/cms/Portals/0/Obesity Views Children R2009Rees.pdf?ver=2010-12-22-121209-040. Accessed 3 Apr 2017.
42. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open. 2016;6:e011458. doi:10.1136/bmjopen-2016-011458.
43. Verboom B, Montgomery P, Bennett S. What factors affect evidence-informed policymaking in public health? Protocol for a systematic review of qualitative evidence using thematic synthesis. Syst Rev. 2016;5. doi:10.1186/s13643-016-0240-6.
44. Newcomer K, El Baradei L, Garcia S. Expectations and Capacity of Performance Measurement in NGOs in the Development Context. Public Adm Dev. 2013;33:62–79. https://search.proquest.com/docview/1735643581?accountid=13042.
45. Thomson DE. The role of funders in Driving Nonprofit Performance Measurement and Use in Strategic Management. Public Perform Manag Rev. 2011;35:54–78. doi:10.2753/PMR1530-9576350103.
46. Minich L, Howe S, Langmeyer D, Corcoran K. Can community change be measured for an outcomes-based initiative? A comparative case study of the success by 6 initiative. Am J Community Psychol. 2006;38:183–90.
47. Kegeles SM, Rebchook GM, Tebbetts S. Challenges and Facilitators to Building Program Evaluation Capacity among Community-Based Organizations. AIDS Educ Prev. 2005;17:284–99. https://search.proquest.com/docview/61300181?accountid=13042.
48. Kang J, Anderson SG, Finnegan D. The evaluation practices of US international NGOs. Dev Pract. 2012;22:317–33. https://search.proquest.com/docview/1735650024?accountid=13042.
49. Akintobi TH, Yancey EM, Daniels P, Mayberry RM, Jacobs D, Berry J. Using evaluability assessment and evaluation capacity-building to strengthen community-based prevention initiatives. J Health Care Poor Underserved. 2012;23:33–48.
50. Hoefer R. Accountability in action?: Program evaluation in nonprofit human service agencies. Nonprofit Manag Leadersh. 2000;11:167–77. https://search.proquest.com/docview/203750114?accountid=13042.
51. Napp D, Gibbs D, Jolly D, Westover B, Uhl G. Evaluation barriers and facilitators among community-based HIV prevention programs. AIDS Educ Prev. 2002;14:38–48. https://search.proquest.com/docview/71861782?accountid=13042.
52. Williams-Gray B. Building Capacity in Nonprofit Human Service Agencies Through Organizational Assessment During the Accreditation Process. J Nonprofit Educ Leadersh. 2016;6. https://search.proquest.com/docview/1796236275?accountid=13042.
53. Mitchell GE, Berlan D. Evaluation and Evaluative Rigor in the Nonprofit Sector. Nonprofit Manag Leadersh. 2016;27:237. https://search.proquest.com/docview/1853309810?accountid=13042.
54. Carman JG, Fredericks KA. Evaluation capacity and nonprofit organizations: Is the glass half-empty or half-full? Am J Eval. 2010;31:84–104.
55. Andrei K, Pope E, Hart A, Quinn LS. The State of Nonprofit Data 2012. 2012. papers2://publication/uuid/7FFEE95D-0054-4393-A722-9D4E7048677C. Accessed 12 Feb 2017.
56. Ruch-Ross H, Keller D, Miller N, Bassewitz J, Melinkovich P. Evaluation of community-based health projects: the healthy tomorrows experience. Pediatrics. 2008;122:e564–72. https://search.proquest.com/docview/69491374?accountid=13042.

57. Big Lottery. An evaluation of Big Lottery Fund's Monitoring and Evaluation Policy for grant holders. 2013. https://www.biglotteryfund.org.uk/research/making-the-most-of-funding/impact-and-outcomes/monitoring-and-evaluation.
58. Morariu J, Athanasiades K, Emery A. State of evaluation 2012: evaluation practice and capacity in the nonprofit sector. 2012. http://www.pointk.org/client_docs/innonet-state-of-evaluation-2012.pdf. Accessed 12 Feb 2017.
59. Hall MH, Phillips SD, Meillat C, Pickering D. Assessing performance: evaluation practices & perspectives in Canada's voluntary sector. 2003.
60. Morariu J, Athanasiades K, Pankaj V, Grodzicki D. State of evaluation 2016: Evaluation Practice and Capacity in the Nonprofit Sector. 2016. https://www.innonet.org/media/2016-State_of_Evaluation.pdf.
61. Morariu J, Reed E. State of Evaluation 2010 : Evaluation Practice and Capacity in the Nonprofit Sector. 2010. http://www.pointk.org/client_docs/innonet-state-of-evaluation-2010.pdf.
62. Bourgeois I, Whynot J, Thériault É. Application of an organizational evaluation capacity self-assessment instrument to different organizations: similarities and lessons learned. Eval Program Plann. 2015;50:47–55. https://search.proquest.com/docview/1673072761?accountid=13042 http://oxfordsfx.hosted.exlibrisgroup.com/oxford?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&genre=article&sid=ProQ:ProQ%3Amedlinesh ell&atitle=Application+of+an+organizational.
63. Moxham C, Boaden R. The impact of performance measurement in the voluntary sector. Int J Oper Prod Manag. 2007;27:826–45. doi:10.1108/01443570710763796.
64. Carnochan S, Samples M, Myers M, Austin MJ. Performance Measurement Challenges in Nonprofit Human Service Organizations. Nonprofit Volunt Sect Q. 2014;43:1014–32.
65. Macmillan R, Ellis A, With P, Kara H, Dayson C, Sanderson E, et al. Building Capabilities in the Voluntary Sector: What the evidence tells us. TSRC Research Report 125 (Birmingham, University of Birmingham). 2014.
https://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/reports/research-report-125-building-capabilities.pdf. Accessed 9 Dec 2017.
66. Building capacity: Research - Big Lottery Fund.
https://www.biglotteryfund.org.uk/research/making-the-most-of-funding/building-capacity. Accessed 14 Dec 2017.
67. Jamieson M. Author's note. Educ Prim Care. 2011;22:123.
68. Millar R, Hall K. Social Return on Investment (SROI) and Performance Measurement. Public Manag Rev. 2012;9037 July 2015:1–19. doi:10.1080/14719037.2012.698857.
69. Pritchett L. The World Bank and public sector management: what next? Int Rev Adm Sci. 2013.
70. Swindle J. The Tough Realities Small NGOs Face When Evaluating Their Efforts: Observations from Guatemala and Mexico. Cons J Sustain Dev. 2011;6:251–63.
71. Arvidson M. Impact and evaluation in the UK third sector: reviewing literature and exploring ideas. TSRC Working Paper 27 (Birmingham, University of Birmingham). 2009.
https://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-27.pdf. Accessed 9 Aug 2016.
72. Arvidson M. Evidence and transparency in the open public services reform: perspectives for the third sector. TSRC Working Paper 117 (Birmingham, University of Birmingham). 2014.

https://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-117.pdf. Accessed 6 Aug 2016.
73. Arvidson M, Lyon F. Social Impact Measurement and Non-profit Organisations: Compliance, Resistance, and Promotion. Voluntas. 2014;25:869–86.
74. Arvidson M, Lyon F, Mackay S, Moro D. The ambitions and challenges of SROI. TSRC Working Paper 49 (Birmingham, University of Birmingham). 2010. http://epapers.bham.ac.uk/788/1/WP49_Ambitions_and_challenges_of_SROI_-_Arvidson%2C_Lyon%2C_McKay%2C_Moro_Nov_2010.pdf. Accessed 20 Oct 2016.
75. Meldrum B, Read P, Harrison C. A Guide to Measuring Social Impact. http://www.learningandwork.org.uk/wp-content/uploads/2017/01/A-guide-to-Measuring-Social-Impact-v2.pdf. Accessed 23 Oct 2016.
76. Rinaldo H. GETTING STARTED IN SOCIAL IMPACT MEASUREMENT - A guide to choosing how to measure social impact. 2010. http://www.socialauditnetwork.org.uk/files/8113/4996/6882/Getting_started_in_social_impact_measurement_-_270212.pdf. Accessed 24 Oct 2016.
77. Millar R, Hall K. Social Return on Investment (SROI) and Performance Measurement. Public Manag Rev. 2013;15:923–41. doi:10.1080/14719037.2012.698857.
78. Hammersley M. Is the evidence-based practice movement doing more good than harm? Reflections on Iain Chalmers' case for research-based policy making and practice. Evid Policy A J Res Debate Pract. 2005;1:85–100.
79. Sackett DL, Rosenberg WMC, Gray J a M, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. Br Med J. 1996;312:71–2.
80. Fraser MW, Galinsky MJ. Steps in Intervention Research: Designing and Developing Social Programs. Res Soc Work Pract. 2010;20:459–66.
81. Montgomery P, Grant S, Hopewell S, Macdonald G, Moher D, Michie S, et al. Protocol for CONSORT-SPI: an extension for social and psychological interventions. Implement Sci. 2013;8:99. doi:10.1186/1748-5908-8-99.
82. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, et al. CONSORT for reporting randomised trials in journal and conference abstracts. Lancet. 2008;371:281–3.
83. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. PLoS Med. 2010;7:e1000217. doi:10.1371/journal.pmed.1000217.
84. Lecy JD, Schmitz HP, Swedlund H. Non-Governmental and Not-for-Profit Organizational Effectiveness: A Modern Synthesis. Voluntas. 2012;23:434–57.
85. Salamon LM, Anheier HK, List R, Toepler S, Sokolowski SW. Civil Society in Comparative Perspective. Glob Civ Soc Dimens Nonprofit Sect. 1999;:3–39.
86. Salamon LM, Sokolowski SW, List R. The Johns Hopkins Comparative Nonprofit Sector Project Global Civil Society An Overview. 2003. http://ccss.jhu.edu/wp-content/uploads/downloads/2011/09/Book_GCSOverview_2003.pdf. Accessed 8 Dec 2017.
87. Lilienfeld SO. Psychological Treatments That Cause Harm. Perspect Psychol Sci. 2007;2:53–70. doi:10.1111/j.1745-6916.2007.00029.x.
88. Petrosino A, Turpin-petrosino C, Hollis-peel ME, Lavenberg JG. 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency. Cochrane Database Syst Rev. 2013;:4–6.
89. Clifford D, Geyne-Rahme F, Mohan J. Variations between organisations and localities in government funding of third-sector activity: Evidence from the national survey of third-sector

organisations in England. Urban Stud. 2013;50:959–76. doi:10.1177/0042098012458550.
90. Mitchell GE. The Construct of Organizational Effectiveness: Perspectives From Leaders of International Nonprofits in the United States. Nonprofit Volunt Sect Q. 2012;42:324–45. doi:10.1177/0899764011434589.
91. Harlock J. From outcomes-based commissioning to social value? Implications for performance managing the third sector. TSRC Working Paper 123 (Birmingham, University of Birmingham). 2014. https://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-123.pdf. Accessed 11 Dec 2017.
92. Ramanadhan S, Crisostomo J, Alexander-Molloy J, Gandelman E, Grullon M, Lora V, et al. Perceptions of evidence-based programs among community-based organizations tackling health disparities: a qualitative study. Health Educ Res. 2012;27:717–28. https://search.proquest.com/docview/1026687544?accountid=13042.

Chapter 4: Barriers and facilitators to implementing evidence-based interventions among third sector organisations: a systematic review[4]

Anders Malthe Bach-Mortensen, MSc[a]

Brittany C.L. Lange, MPH[a]

Paul Montgomery, DPhil, MSc, DipSW[b]

[a] University of Oxford, Department of Social Policy and Intervention; Barnett House, 32 Wellington Square; Oxford; UK

[b] University of Birmingham, Department of Social Policy, Sociology and Criminology, Birmingham, B15 2TT, UK.

Author contributions: The research design and protocol were developed by ABM and PM with helpful feedback from BL. ABM designed and conducted the electronic search. All identified studies were double-screened and data were extracted independently by ABM and BL. Quality appraisal was also conducted independently by ABM and BL with any disagreements resolved through discussion until consensus was reached. The synthesis was conducted by ABM with feedback from BL and PM. ABM drafted the manuscript with input from PM and BL. All authors read and approved the final manuscript.

[4] This paper has been published in Implementation Science.

Abstract

Background: The third sector is becoming a growing provider of public, social, and health services. However, there is little evidence on the effectiveness of third sector organisations (TSOs), and their capacity to implement evidence-based interventions (EBIs). Understanding implementation aspects of service delivery remains an important issue in clinical practice, but is poorly understood in the context of TSOs. This is problematic, since implementation issues are known to be critical for effective intervention outcomes.

Objectives: To identify and synthesise existing research on what barriers and facilitators influence the implementation process of TSOs delivering EBIs.

Methods: This review is reported according to PRISMA guidelines and was pre-registered in PROSPERO. Key databases were searched using relevant terms, experts in the field were contacted, and websites were reviewed. All identified studies were double-screened and data were extracted independently by two authors. Included studies were synthesised using thematic analysis and were quality appraised.

Results: Thirty-one studies were included, most of which were conducted in North America. The thematic synthesis identified *resource limitations*, in particular staff and finance, to be the most reported barrier to TSOs implementing EBIs. *Organisational culture*, including factors such as alignment between the mission of the TSO and EBI, and *support/prioritisation* of the implementation process, were the most reported facilitators. These findings generalise across the included studies and are robust to study quality assessment.

Conclusions: While it is often assumed that good outcomes follow when implementing interventions that have been developed and tested according to best practice, little attention has

been paid to how EBIs are best transported, contextualised, and implemented by third sector providers. This systematic review found that TSOs faced considerable challenges in implementing EBIs, which were primarily a lack of support and expertise, and unclear/insufficient guidelines on how to adapt EBIs to different populations. To address these challenges, it is important to engage with central stakeholders, such as funders, researchers, policymakers, and practitioners to discuss how these needs can be met.

Registration: PROSPERO: CRD42017073090

Key-words: Evidence-based intervention; EBI, evidence-based practice; EBP, third sector organisation; TSO; implementation; barriers; facilitators; enablers; obstructers; charities; non-profits; CBO; NGO.

Background

The third sector is expanding and becoming a growing provider of public, social and health services in many high-income countries [1–4]. However, little research demonstrates evidence on the *effectiveness* and *impact* of third sector service deliveries [2, 3, 5, 6].

While scholars struggle to agree on a universally applicable definition of the third sector [7, 8], most follow Anheier and Salamon (1997), who point to five characterising traits of third sector organisations (TSOs); i.e. that they are: (1) formally structured, (2) privately owned and independent from the government, (3) non-profit distributing, (4) self-governing, and (5) benefitting from voluntary activities [5, 9, 10]. However, the use of terminology is variable and terms such as 'non-profits', 'NGOs', 'community-based organisations' (CBOs), 'charities', 'voluntary organisations' etc. are often used synonymously. To avoid confusion on the arbitrary distinction between these terms, we will adopt the more commonly applied term 'third sector organisation' to denote any such organisation [7, 11].

The third sector is often considered to entail distinct features and characteristics in its service delivery compared to public and private sector bodies [2–5, 12]. For instance, TSOs are thought by some to be better at connecting with hard-to-reach populations, while also being driven by more altruistic values [5, 13, 14]. These perceived traits often appear in the discourse of policymakers, who continue to emphasise the growing importance of the third sector, especially in the context of alleviating social problems [3, 15–18]. Yet, despite the political willingness to increase the role of the third sector in public service provisions, there is little research to demonstrate evidence on the *capacity* and *capability* of TSOs to successfully adhere to evidence-based practice (EBP) and to implement evidence-based interventions (EBIs) [4, 19, 20].

Adherence to EBP implies having established feedback mechanisms between services and outcomes, and the inclusion of stakeholders and the best available evidence in decision-making processes [21–24]. In enabling TSOs to become more evidence-based, there has been a growing emphasis on the development of 'evidence hubs' (e.g. What Works Network [25], Project Oracle [26], Blueprints for Healthy Youth Development [27], and Diffusion of Evidence Based Interventions (DEBIs) [28]), which are efforts to develop disseminating databases of interventions and programmes that are 'proven to work'. A set of proven programmes could enable service providers (both private, public, and third sector) to adopt and implement programmes that are supported by sound scientific evidence with the highest potential for effectiveness. However, while TSOs are known to deliver EBIs in the context of, for example, HIV [29] and addiction [30], there is little systematic research providing a general overview on the types of EBIs implemented by the third sector.

Whilst the principles of EBP are gradually becoming an integrated part of common and accepted practice in social work and clinical settings, practitioners still experience substantial barriers to implementing EBIs [31, 32]. Implementation is often understood in terms of *implementation fidelity,* which can be defined as "the degree to which interventions are implemented according to the design." [33]. Implementation fidelity is inevitably linked to the concept of *adaptation,* which can be understood as "changes made in the original program during implementation (program modification, reinvention)" [34]. There is an ongoing debate on how practitioners should balance fidelity and adaptation when implementing EBIs, which will be discussed later in this review. Further implementation aspects include adherence to protocol, dose of treatment, participant responsiveness to intervention, and implementation quality [33, 35].

Past research has shown that most intervention studies fail to report adequately on implementation aspects such as fidelity and adaptation, which represent an important blind-spot

in understanding 'true' intervention effects and in conducting meaningful replications [34, 36–38]. The failure to understand aspects related to implementation introduces the risk of overlooking type iii errors ('implementation failure') [39, 40], i.e. failure to implement an intervention as intended [41]. Overlooking this issue has great implications for policy and practice, as it may lead to false inferences about the effectiveness of interventions and programmes.

Thus, without understanding the *implementation aspects* of third sector service deliveries, it is difficult to assess their potential to substitute for public sector provision of social and health services. Just as relatively little is known about implementation aspects of EBP in clinical settings [31], there is even more limited research on the capacity of TSOs to implement EBIs. This failure to understand aspects of the implementation ability of TSOs is worrisome, in that such a research gap questions the potential of TSOs to become evidence-based service providers. As the role of the third sector is becoming more salient in the delivery of public and social services, it is critical to ensure that such organisations are willing and able to implement effective and safe services supported by the appropriate evidence-base; especially considering that interventions have the potential to do harm to service-users [42–44]. Further, most TSOs in, for example, the UK work around 'social services' and target vulnerable population groups, such as disabled people, children, and the elderly [45], which arguably warrant the use of EBP.

What has been done?

A range of single studies have investigated the experiences and attitudes of TSOs to implementing EBIs and adhering to EBP [29, 30, 46–48], investigating topics such as barriers and facilitators for TSOs to implementing EBIs [30], perceived needs of TSOs adhering to EBP [49], and the attitudes of third sector practitioners to adopting EBIs [49, 50]. Identified barriers to the implementation of EBIs by TSOs tend to involve factors related to organisational culture,

such as staff resistance and organisational setting [30, 51], and factors related to the lack of resources [47]. Facilitators include having established affiliations with research institutions and employing skilled staff [46, 47]. An overarching theme of these studies seem to revolve around the notion that TSOs struggle to implement EBIs and experience serious capacity issues in becoming evidence-based providers.

To date, there has been no systematic attempt to aggregate and analyse existing research on the implementation ability of TSOs delivering EBIs. This constitutes a significant knowledge gap, given that research continues to demonstrate that implementation aspects are critical to the effectiveness of service deliveries [23, 36, 38, 52]. Also, the utility of the increasingly popular 'evidence hubs' and 'blueprints' relies on the ability of practitioners to *implement* the EBIs according to best practice.

Objectives

To utilise the full potential of TSOs in the delivery of social and health services, it is crucial to understand what factors influence their implementation process, so that the commissioning and regulation criteria can ensure that delivery is conducted following best practice. The main objective of this review is to aggregate existing research investigating practitioner-identified factors affecting the implementation process of TSOs that deliver EBIs. The focus of the study is captured by the following question:

- What barriers and facilitators influence the implementation process of third sector organisations delivering evidence-based interventions and programmes?

Methods

To meet these research objectives, a systematic review was conducted following PRISMA guidelines (see appendix B1 for completed PRISMA checklist) [53]. The protocol was reported according to PRISMA-P guidelines [54] and was pre-registered in PROSPERO (CRD42017073090).

Search strategy

The search strategy was designed to be exhaustive of the existing literature on studies investigating barriers to, and facilitators of, TSOs delivering EBIs (see appendix B2 for search terms). A body of research addressing this topic was identified prior to the search through Google Scholar and subsequent reference checking. The bibliography of the pre-identified literature was hand-searched to identify further studies.

The following databases were searched during the systematic review using text words: ABI/INFORM Global, Applied Social Sciences Index & Abstracts (ASSIA), International Bibliography of the Social Sciences (IBSS), MEDLINE®, PAIS Index, Policy File Index, Social Services Abstracts, Worldwide Political Science Abstracts, Social Care Online, SCOPUS, and Open Grey.

The search was revised until it was sensitive enough to capture at least all the pre-identified studies. The final body of included studies was hand-searched for additional references that may not have been captured by the search. Experts in the field were contacted and websites of key organisations reviewed.

Selection criteria

To be eligible for inclusion, studies had to be primary research or systematic reviews

investigating the perspectives and/or experiences of third sector practitioners (e.g. managers, directors or service providers) with regards to the implementation/adoption of EBIs. We considered all studies that investigated the process of delivering an EBI. This process might include, but is not limited to, aspects such as *fidelity* to intervention protocol, and whether *adaptations* were made in the delivery of the intervention. We define EBIs as "interventions which have been tested and validated according to the principles of evidence-based practice". EBIs may be implemented in the context of, for example, social care, healthcare, education, child services, or mental health services, but other contexts of implementation were also considered for inclusion. We define barriers as "any factors that obstruct the capacity for third sector organisations to implement evidence-based interventions". We define facilitators as "the factors that enable the implementation of evidence-based interventions".

All research designs were eligible for inclusion. To be eligible for inclusion, studies had to investigate factors operating as barriers and/or facilitators to the implementation of evidence-based programmes and interventions by TSOs, but this did not need to be the focus of the studies. If it was unclear whether the samples included TSOs, the authors of the studies were contacted for clarification. Articles were screened at the title and abstract level independently by both ABM and BL using the Rayyan systematic review software [55].

Data extraction

All data were independently double-extracted by ABM and BL with the following information being retrieved:

- Publication year and author
- Study aim
- Methods (study design, data collection methods, data analysis, and inter-rater reliability)
- Population (type of organisations, area of work, and sample size)

- Types of EBIs implemented
- Results (barriers and facilitators in implementing EBIs)
- Discussion (suggestions for future research and policy)

For qualitative studies, interview and focus-group quotes were double-extracted in separate documents (see appendix B3 and B4). These quotes were utilised to extract factors not captured by the individual studies, and to evaluate the reliability of the subsequently constructed themes. After finishing the double-extraction, both reviewers (BL and ABM) met in person to review all identified factors, which were discussed until consensus was reached. Upon reaching consensus, the synthesis was initiated.

Data synthesis

To analyse the included studies, a thematic analysis was conducted, in line with best practice when aggregating data from different types of research [56–59]. Specifically, all identified factors were identified and organised into barriers and facilitators and counted by frequency. The identified factors were then categorised following thematic analysis [59], thus enabling the synthesis to account for the arbitrary difference of factors revolving around the same underlying problem. Identified factors were only counted once per study, except for studies which identified factors specific to different subgroups (e.g. according to different organisational cases [60] or by different types of EBIs [30]). To the best of our knowledge, all included studies investigated unique samples.

The construction of themes was done inductively and entirely according to the factors identified in the data extraction. ABM conducted the full thematic analysis, which was reviewed by BL and PM on an iterative basis. All modifications were made through discussion until consensus was reached. Additionally, we constructed two tables (Table 5 and 6) following Rees et al (2009) [61]

to provide an overview of how the individual studies contributed to the construction of the identified themes. A study was considered to contribute to a theme if it identified at least one factor part of that theme. This enabled an assessment of whether certain studies were over- or under-represented in the thematic framework [56]. Further, it allowed for assessing the reliability of the identified themes based on an overall judgement of the quality of the studies that contributed to the individual themes. A sensitivity analysis was conducted excluding studies of low quality to test the robustness of findings.

Quality appraisal

To ensure transparency, all included studies were subject to best practice quality appraisal. For qualitative studies, a modified version of the Joanna Briggs Institute and CASP checklist was applied (see appendix B5) [62]. For the appraisal of survey studies, a modified version of the AXIS checklist was employed (see appendix B6). For mixed-methods studies, the appropriate quality appraisal tool was decided according to the type of method employed by the study to identify barriers and facilitators to implementation. These tools allowed for an overall assessment of key biases of the included studies and the final quality ratings were subsequently utilised to assess the reliability of the identified factors. All appraisals were conducted independently by ABM and BL with any disagreements resolved through discussion until consensus was reached.

Results

2,654 articles were identified through the database searches of which 1,850 remained after removal of duplicates. 1,722 studies were excluded based on screening titles and abstracts. 128 studies were reviewed in full text, in which studies were excluded for not being primary research (n=32), not being TSOs (n=31), including a mixed sample of organisations (n=18), not investigating barriers and facilitators (n=6), not focusing on EBIs (n=5), not focusing on the

implementation process (n=4), and not reporting sufficiently on the results (n=1). Thirty-one studies were included for the thematic synthesis, which were all identified through the database searches. No additional studies were retrieved via searches of websites or by reviewing the references lists of the included studies. All studies suggested by the contacted experts or identified via websites that were eligible for inclusion had already been identified by the database searches.

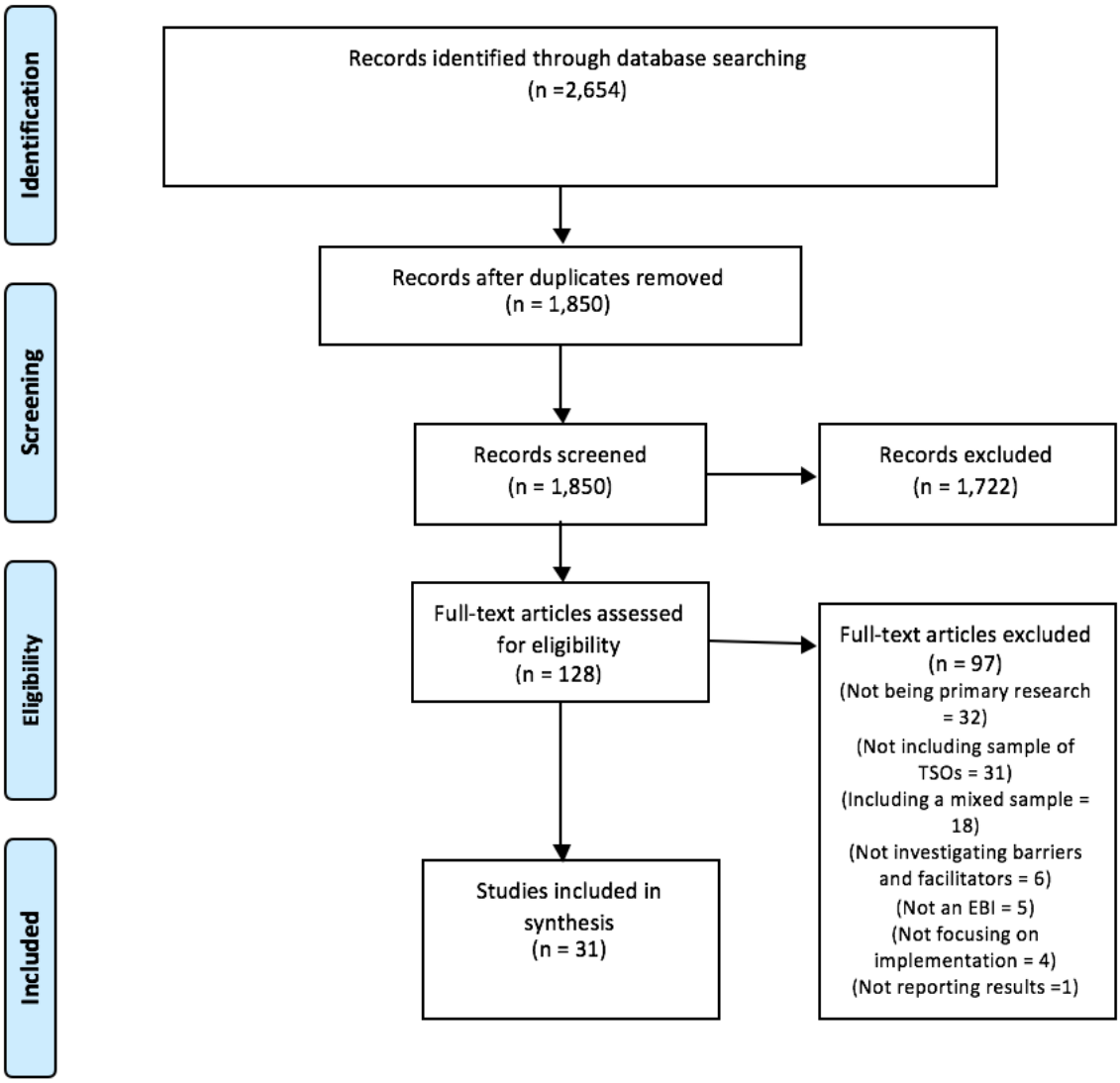


Figure 1: PRISMA diagram

Characteristics of included studies

All included studies were published after 2009 in peer-reviewed journals, except for two doctoral dissertations [63, 64]. Of the 31 included studies, 26 employed qualitative methods [19, 29, 30, 47, 48, 50, 60, 64–79], and 5 were cross-sectional [49, 80–83].

Sample size of included studies

In the research employing qualitative methods, 4 studies included a sample of 1-10 practitioners [60, 64, 74, 77], 16 studies included a sample of 11-50 practitioners [29, 47, 48, 63, 65, 67–69, 71–73, 75, 76, 79, 84, 85], 2 studies included a sample of 51-75 practitioners [19, 70], and 3 studies investigated more than 100 participants [30, 50, 66]. One article did not provide clear information on its sample [78]. For the quantitative articles, 1 study investigated a sample of fewer than 10 TSOs [80], 1 study included 82 TSOs [49], 1 study included a sample of 100 TSOs [49], 1 study included 112 practitioners from 41 NGOs [82], and the last study investigated 510 staff members and 296 directors [81].

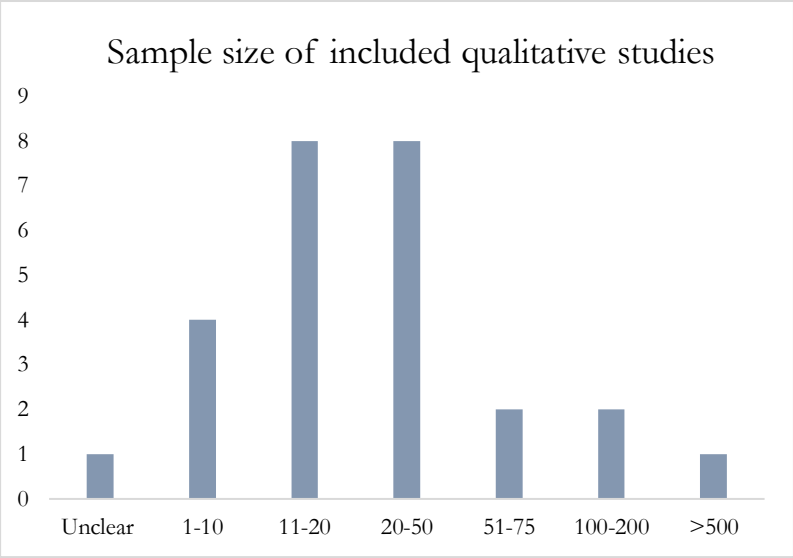


Figure 2: Sample size of included qualitative studies

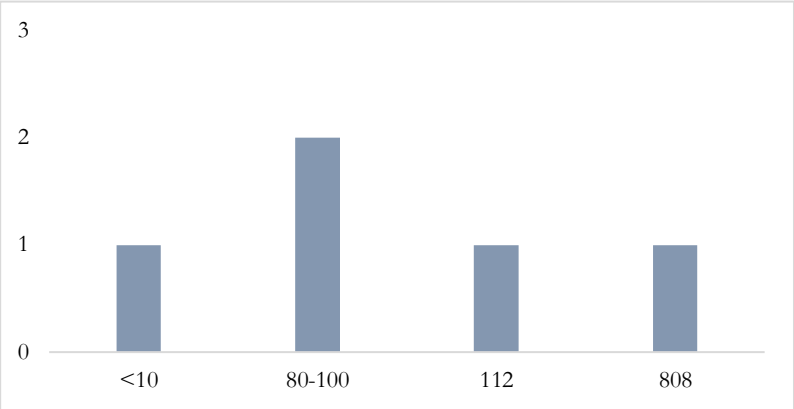


Figure 3: Sample size of included quantitative studies

Types of organisations

The majority of the included research was conducted in the United States (US) (28/31) [19, 29, 30, 47, 48, 50, 60, 63–66, 68–81, 83–85], with three studies being conducted in Chile [67], Puerto Rico [49], and Australia [82]. Most studies (23/31) referred to the included organisations as "community-based organisations" (CBOs) [19, 29, 30, 47–50, 60, 63, 64, 66, 68, 70–72, 74–76, 78, 80, 81, 83, 85], with five studies using the term "non-profits" [65, 67, 69, 73, 77], two referring to its included sample as "churches" [79, 84] (in the US, churches are considered non-profits), and one study denoting its included organisations as NGOs [82].

Types of EBIs

All included studies investigated EBIs in the context of health and social outcomes, including addiction [30, 63, 66, 83, 86, 87], HIV [19, 29, 48, 64, 68, 71, 75, 76], exercise [65, 70, 73, 74], social work [67], cancer education [85], parenting [77], nutrition [79, 84], and with a number of studies focusing on mixed types of health EBIs [47, 49, 69, 72, 82, 85] (see Table 1). Nine studies focused on factors influencing the implementation of interventions from the Diffusion of Evidence based interventions (DEBI) programme by the Center for Disease Control and Prevention (CDC) [19, 29, 48, 64, 68, 71, 72, 76, 78], and six studies investigated EBIs designed by the Center for Substance Abuse Treatment (SAMHSA/CSAT) [30, 63, 66, 83, 86, 87].

Table 1: Types of EBIs investigated in included studies.

Types of EBIs	Cited in
Addiction-related EBIs	
– Mixed types of EBIs supported by SAMHSA/CSAT	(Dippolito et al., 2015; Kegeles et al., 2012; Lundgren et al., 2011)
– Adolescent Community Reinforcement Approach	(Amodeo et al., 2011; Hunter, Han, Slaughter, Godley, & Garner, 2017)
– **Motivational Interviewing**	(Amodeo et al., 2011)
– Assertive Community Treatment	(Amodeo et al., 2011)
– Cognitive Behavioural Therapy	(Amodeo et al., 2011)
– Integrated Dual Disorder Treatment	(Maharaj, 2010)
HIV related EBIs	
– Mixed types of DEBIs	(Collins et al., 2010; Gandelman & Dolcini, 2012; Margaret Dolcini et al., 2010; Owczarzak, 2012; Owczarzak & Dickson-Gomez, 2011; Veniegas, Kao, & Rosales, 2009)
– The Mpowerment Project	(Kegeles, Rebchook, Tebbetts, & Arnold, 2015)
– "The Modelo de Intervención Psicomédica"	(Pemberton, 2012)
Mixed health-related EBIs	
– EBIs for underserved populations	(Payan et al., 2017; Ramanadhan et al., 2012)
– Mixed health EBIs	(Kimber, Barwick, & Fearing, 2012)
– EBIs for youth having experienced abuse	(Thomas, Zimmer-Gembeck, & Chaffin, 2014)
– Mixed cancer education EBIs	(Vanderpool et al., 2011)
– Unspecified	(Martínez et al., 2014)
Exercise EBIs	
– Enhance®Fitness	(Belza et al., 2015; Petrescu-Prahova, Belza, Kohn, & Miyawaki, 2016)
– Home-based exercise program for breast cancer survivors	(Pinto, Waldemore, & Rosen, 2015)
– Mixed exercise EBIs	(Lattimore et al., 2010)
Sexual education	
– Becoming a Responsible Teen (BART)	(Demby et al., 2014)
– Teen pregnancy-prevention EBIs	(House, Tevendale, & Martinez-Garcia, 2017)
– "Cuicidate"	(Feutz & Andresen, 2013)
Nutrition focused EBIs	
– "Body and Soul"	(Allicock et al., 2013)
– Mixed EBIs on increasing vegetable and fruit consumption	(Honeycutt, Carvalho, Glanz, Daniel, & Kegler, 2012)
Social work	
– Unspecified	(Flores, Naranjo, & Hein, 2016)
Parenting EBIs	
– "Together Facing the Challenge"	(Murray, Culver, Farmer, Jackson, & Rixon, 2014)

Quality of studies

Overall, 29 of the included studies were rated to be of either high or medium quality, with only three studies being rated as low quality. Of the 26 included qualitative studies (see appendix B8 3), nine were rated to be of high quality, fifteen of medium quality, and two of low quality. Of the five included survey studies (see appendix B9), four were rated to be of medium quality, and one study to be of low quality.

Barriers and facilitators

The synthesis identified 80 unique factors operating as barriers across 31 studies and 57 factors operating as facilitators over 24 studies. Table 2 illustrates the five most reported factors operating as barriers and facilitators to implementation of EBIs. The most reported barriers were related to recruitment and retention of service-users (14/31), problems in adapting EBIs (13/31), lack of financial resources (13/31), lack of staff resources (11/31), and implementation difficulty (9/31).

The most reported facilitating factors were related to whether the EBI matched with the mission of the TSO (9/24), flexibility for TSOs to implement the EBI (9/24), perceived effectiveness of the EBI (8/24), organisational support and prioritisation of the EBI, and supportive leadership (5/24).

Table 2: Top 5 most reported factors operating as barriers and facilitators.

Top 5 cited barriers and facilitators	
Barriers	**Facilitators**
Recruitment/retention issues 14/31	EBI matches well with mission of TSO 9/24
Problems in adapting the EBI 13/31	Flexibility regarding the implementation of interventions 9/24
Lack of financial resources 13/31	Perceived effectiveness of EBI 8/24
Lack of staff resources/high staff turnover 11/31	Organisational support/prioritization of EBI 6/24
Implementation difficulty/fidelity issues 9/32	Supportive leadership 5/24

Thematic analysis

The main results of the synthesis are illustrated in Table 3 and 4, which display the prevalence of identified factors organised according to their underlying theme. All identified factors were given a reliability rating to reflect how consistently each factor was reported on and the quality of studies identifying it (see Ratings Key).

The most prevalent themes impeding implementation revolved around resources (e.g. lack of time, finances, and staff), followed by client and community factors (e.g. recruitment/retention issues), delivery capability (e.g. lack of expertise), and organisational culture (e.g. conflict with EBI and TSO mission). Other significant themes included challenges around adapting the EBI (e.g. not knowing how to adapt the intervention to the target population), and lack of training. Less significant themes involved monitoring issues, intervention-specific problems (e.g. transportation and legal barriers), commissioning requirements (e.g. paperwork), and collaboration issues.

Table 3: Themes of factors operating as barriers.

Themes	Times cited	Reliability of findings
Resources	**51**	
Lack of financial resources	13	High reliability
Lack of staff resources/high staff turnover	11	High reliability
Lack of time	8	High reliability
Insufficient space for all services	5	High reliability
Resources (unspecified)	5	High reliability
Lack of resources for training	3	High reliability
Lack of scientific resources	3	Medium reliability
Lack of technical resources	3	Medium reliability
Client and community factors	**37**	
Recruitment/retention issues	14	High reliability
Client resistance and non-participation	5	High reliability
Difficult population	5	High reliability
Lack of community support/community resistance	4	High reliability
Client receptiveness and commitment	3	High reliability
Client (individuals and family) attendance is poor	2	High reliability
Community resources (including substance abuse treatment) are lacking	2	High reliability
Stigma	1	Low reliability

Anonymity issues	1	Low reliability
Delivery capability	**33**	
Implementation difficulty/fidelity issues.	9	High reliability
Lack of expertise/experience	7	High reliability
Lack of administrative infrastructure	3	High reliability
Cultural/language barriers	3	High reliability
Competing responsibilities	2	Medium reliability
Supervision issues	2	Medium reliability
Program needs	2	Medium reliability
Delivering the EBI in a rural area	1	Medium reliability
Scheduling challenges	2	Low reliability
Strategic planning	1	Low reliability
Balancing needs between of the CBO and research	1	Low reliability
Organisational culture	**32**	
Conflict with EBI and organizational identify/mission/culture	6	High reliability
Staff resistance	5	High reliability
Lack of prioritization of the EBI/org support	3	High reliability
Lack of preparation of staff to become evidence based providers	3	High reliability
Lack of leadership	3	Medium reliability
Incomplete buy-in from organization	2	Medium reliability
Resistance to change (keeping status quo)	2	Medium reliability
Change in staff and leadership	2	Medium reliability
Lack of EBI champions	1	Medium reliability
Lack of belief in the efficacy of the intervention	1	Medium reliability
Unclear mission of the organisation	1	Medium reliability
Higher level of stress	1	Medium reliability
Low cohesion in the organization	1	Medium reliability
EBP devaluing existing practices by the TSO	1	Low reliability
Adaptation issues	**22**	
Problems in adapting the EBI	13	High reliability
Lack of fit between EBI and target population of TSO	7	High reliability
Conflict between the EBI and expert knowledge of the provider	1	Low reliability
Not knowing how to communicate adaptation issues	1	Low reliability
Training capacity issues	**19**	
Staff not trained well enough/ Training provided for the EBI not adequate	8	High reliability
Lack of EBI training	4	High reliability
Variance in staff training and perspectives	2	Medium reliability
Lack of training available	1	Low reliability
Training new staff	1	Low reliability
Programs were not packaged for training	1	Low reliability
Conducting community needs assessments	1	Low reliability
Proposal development	1	Low reliability
Monitoring issues	**10**	
Not having monitoring practices in place to evidence effectiveness.	4	High reliability
Data management issues	2	Medium reliability
Interoperability challenges between systems	1	Medium reliability
Issues around evaluation and monitoring of implementation	3	Low reliability

Intervention specific barriers	**10**	
Transportation	3	High reliability
State education law in conflict with EBI	1	Low reliability
State law prohibits collection of certain data	1	Low reliability
State law requires approval by school board and parental review committee	1	Low reliability
Assessing light versus moderate activity	1	Low reliability
Incorporating physical activity into the participants' lifestyle	1	Low reliability
Geographically dispersed offices made consistent implementation difficult	1	Low reliability
Logistical issues	1	Low reliability
Commissioning requirements	**8**	
Certification process was burdensome/time consuming	3	High reliability
Lack of support by funder	2	Medium reliability
Prohibition on adaptation by funders.	1	Medium reliability
Funder demand for major modifications to EBI	1	Low reliability
Funder rejection of modifications to EBI	1	Low reliability
Collaboration issues	**6**	
Hard to establish collaboration agreements	2	Medium reliability
Lack of strategies for community mobilization	1	Low reliability
Lack of integration of services	1	Low reliability
Hard to find potential academic partners	1	Low reliability
Assistance in strategies for the development of EBIs	1	Low reliability
Others	**3**	
Competition with other afterschool activities	1	Low reliability
Development and analysis of public policies	1	Low reliability
Inability to fund program partners because of local agency policies	1	Low reliability

Key for Reliability Ratings*

High reliability	The identified factor is consistently supported by several studies of medium quality and from one high-quality study, or the study is supported by at least two high-quality studies.
Medium reliability	The identified factor is supported from several medium-quality studies, or the factor is identified from at least one high-quality study.
Low reliability	The identified factor is supported by several studies of low quality and/or single studies of medium quality.

*It should be noted that confidence in findings do not relate to the *generalisability* of the findings.

The thematic synthesis of factors operating as facilitators identified the main category to be organisational culture, which involved factors such as organisational support and alignment between the organisational mission and the EBI. Other significant themes included delivery capability (e.g. staff expertise), accountability (e.g. perceived effectiveness of the EBI), and adaptation (e.g. flexibility and match with target population). Less reported factors involved collaboration (e.g. with academic and experienced partners), client/community factors (e.g.

community support), funders (e.g. continuous support), resources, and intervention-specific factors (e.g. availability of manuals).

Table 4: Themes of factors operating as facilitators.

Themes	Times cited	Reliability of findings
Organisational culture	**40**	
EBI matches well with mission of TSO	9	High reliability
Organisational support/prioritization for EBI	6	High reliability
Champions	4	High reliability
Supportive leadership	5	Medium reliability
Inclusive organisational team/culture	4	Medium reliability
Staff motivation to deliver EBI	3	Medium reliability
Organisational capacity	3	Medium reliability
Strategic planning	2	Medium reliability
Organisational stability	1	Medium reliability
Internal flexibility of resources	1	Low reliability
Involving staff and volunteers	1	Low reliability
Strength of planning committee	1	Low reliability
Delivery capability	**18**	
Staff expertise	4	High reliability
Receiving intervention training	4	High reliability
Staff experience with delivering EBI	4	Medium reliability
Administrative infrastructure and capacity	4	Medium reliability
Continuous training	1	Low reliability
Clear staff roles in implementing the EBI	1	Low reliability
Accountability	**15**	
Perceived effectiveness of EBI	8	High reliability
Continuous evaluation of EBI	3	High reliability
Novelty of EF	1	Medium reliability
Relative advantage of EBI	1	Low reliability
Cost effective	1	Low reliability
Using EBI will improve professional practice	1	Low reliability
Adaptation	**13**	
Flexibility regarding the implementation of interventions	9	High reliability
Match with the target population	4	High reliability
Collaboration/external factors	**10**	
Invitation to partner with another organisation to offer EF	1	Medium reliability
Working with partner with experience delivering the EBI	2	Low reliability
Working with experienced partners	2	Low reliability
Working with partners with consistent access to target population	1	Low reliability
Partnerships (unspecified)	1	Low reliability
Political support	1	Low reliability
Coordination and communication with other agencies	1	Low reliability
External support	1	Low reliability

Clients/community factors	**9**	
Use of incentives	2	Medium reliability
Access to target population through existing programs	2	Low reliability
Experience working with population	2	Low reliability
Ability to recruit and retain IDU clients	2	Low reliability
Client participation in goal setting and feedback.	1	Low reliability
Funders	**7**	
Access and availability of training	2	Medium reliability
Continuous support in delivering the EBI.	2	Medium reliability
Using EBI that is recognised by funder	1	Medium reliability
Funding stability	1	Low reliability
Flexibility of prioritisation of resources	1	Low reliability
Other	**7**	
Logistics	1	Low reliability
Established national organization	1	Low reliability
Communication	1	Low reliability
HIV testing success	1	Low reliability
No. of clinicians certified/employed at grant end	1	Low reliability
No. of supervisors certified/employed at grant end	1	Low reliability
No. of youth served during grant period	1	Low reliability
Resources	**4**	
Availability of technical assistance	2	Medium reliability
Financial support	1	Medium reliability
Supplying needed resources	1	Low reliability
Intervention specific factors	**3**	
Easy to use	1	Low reliability
Availability of manuals	1	Low reliability
Implementation is rewarding	1	Low reliability

Key for Reliability Ratings*

High reliability	The identified factor is consistently supported by several studies of medium quality and from one high-quality study, or the study is supported by at least two high-quality studies.
Medium reliability	The identified factor is supported from several medium-quality studies, or the factor is identified from at least one high-quality study.
Low reliability	The identified factor is supported by several studies of low quality and/or single studies of medium quality.

*It should be noted that confidence in findings do not relate to the *generalisability* of the findings.

Robustness of findings

To assess the consistency of the identified factors and the quality of the studies reporting them, all factors were given a reliability rating (see key for reliability rating). Further, to ensure that the findings were not driven by the quality of the included research, we re-conducted the thematic

analysis excluding the studies of low quality [77, 78, 80], which did not result in any important changes to the thematic categories or the ranking of top cited barriers and facilitators.

Table 5: Overview of the contribution of individual studies on the identified themes of barriers.

Study	Study methodology	Factors related to lack of resources	Factors related to client and community factors	Factors related to delivery capability	Factors related to organisational culture	Factors related to adaptation issues	Factors related to training capacity issues	Factors related to monitoring issues	Intervention specific issues	Factors related to commissioning requirements	Factors related to collaboration issues
Allicock et al. 2012	Qualitative	X	X	X							
Amodeo et al. 2011	Qualitative	X	X	X	X	X	X	X	X	X	X
Belza et al. 2014	Qualitative	X		X	X						
Collins et al. 2010	Qualitative	X	X								
D'Ippolito et al. 2015	Qualitative	X					X				
Demby et al. 2014	Qualitative	X	X	X					X		
Dolcini et al. 2010	Qualitative	X	X	X	X	X					
Feutz et al. 2013	Qualitative	X	X	X							
Flores et al. 2016	Qualitative					X					
Gandelman et al. 2012	Qualitative		X	X		X					
Honeycutt et al. 2012	Qualitative	X			X						
Kegeles et al. 2015	Qualitative	X		X	X	X	X	X		X	
Kimber et al. 2012	Qualitative	X			X						
Lattimore et al. 2010	Qualitative	X	X	X		X			X		
Lundgren et al. 2011	Qualitative	X		X	X	X	X				
Maharaj 2010	Qualitative	X		X	X						
Murray et al. 2014	Qualitative	X		X	X			X	X		
Owczarzak et al. 2011	Qualitative	X	X	X	X	X	X			X	

Owczarzak 2012	Qualitative	X		X	X	X		X			
Payan et al. 2017	Qualitative	X	X	X				X			
Pemberton et al. 2012	Qualitative	X	X	X	X	X				X	
Petrescu-Prahova et al. 2016	Qualitative	X	X	X	X		X				
Pinto et al. 2015	Qualitative	X		X							
Ramanadhan et al. 2012	Qualitative	X		X	X	X		X		X	
Vanderpool et al. 2011	Qualitative	X	X			X					
Veniegas et al. 2009	Qualitative	X	X	X		X	X	X		X	
House et al 2017	Cross-sectional		X				X				
Hunter et al 2017	Cross-sectional			X	X						
Lundgren et al 2012	Cross-sectional	X		X	X	X	X				
Martinez et al 2014	Cross-sectional	X		X			X	X			X
Thomas et al 2014	Cross-sectional	X		X	X	X	X				
Total		**27**	**15**	**24**	**17**	**15**	**10**	**8**	**4**	**6**	**2**

Please note that each study can contribute with multiple factors to the same theme. Shading indicates studies that were rated to be of low quality.

Table 6: Overview of the contribution of individual studies on the identified themes of facilitators.

Study	**Study methodology**	Factors related to organisational culture	Factors related to delivery capability	Factors related to accountability	Factors related to adaptation	Factors related to collaboration	Factors related to client/community factors	Factors related to funders	Factors related to resources	Intervention specific issues
Belza et al. 2014	Qualitative	X		X	X	X			X	
D'Ippolito et al. 2015	Qualitative	X	X					X		
Demby et al. 2014	Qualitative	X	X				X			
Dolcini et al. 2010	Qualitative				X		X			

Feutz et al. 2013	Qualitative		X				X			
Flores et al. 2016	Qualitative				X					
Honeycutt et al. 2012	Qualitative	X	X					X		
Kegeles et al. 2015	Qualitative	X	X	X						
Kimber et al. 2012	Qualitative	X		X					X	
Lundgren et al. 2011	Qualitative				X					
Maharaj 2010	Qualitative	X	X		X	X	X			
Murray et al. 2014	Qualitative	X	X		X					
Owczarzak et al. 2011	Qualitative	X			X	X		X		
Owczarzak 2012	Qualitative	X		X						
Pemberton et al. 2012	Qualitative	X	X	X			X	X		
Petrescu-Prahova et al. 2016	Qualitative	X	X							
Pinto et al. 2015	Qualitative	X								
Ramanadhan et al. 2012	Qualitative			X	X			X	X	
Vanderpool et al. 2011	Qualitative				X					
Veniegas et al. 2009	Qualitative	X								
House et al 2017	Cross-sectional					X				
Hunter et al 2017	Cross-sectional	X		X		X		X		
Martinez et al 2014	Cross-sectional	X		X	X	X			X	X
Thomas et al 2014	Cross-sectional	X	X	X				X		X
Total		**17**	**10**	**9**	**10**	**6**	**5**	**7**	**4**	**2**

Please note that each study can contribute with multiple factors to the same theme. Shading indicates studies that were rated to be of low quality.

Tables 5 and 6 demonstrate how each study contributed to the identified themes. For both barriers and facilitators, the most reported themes were also the ones most representative of the included research. However, there were several inconsistencies between the representativeness of

the themes and how often certain categories were reported on. For example, the second most reported category 'factors related to client and community issues' was identified in fewer studies than the third and fourth most reported categories. Further, the third most reported category 'factor related to delivery capability' were more representative of the included research than the second most reported category. Similarly, for facilitators the second, third, and fourth most reported categories represented roughly the same amount of studies.

There was no clear pattern to suggest that study quality or methodology affected what thematic categories to which the included studies contributed.

Recommendations for policy and practice

Additional to the analysis of factors operating as barriers and facilitators, we also aggregated the recommendations posed by the included research. The main identified themes were split between recommendations targeted at: (1) funders and collaborators; (2) research and practice; and, (3) TSO practitioners. The full results can be found in Table 7, and less salient recommendation factors can be found in appendix B7.

Table 7: Main recommendation categories of the included research.

Main Themes	Elaboration
Recommendations for funders and collaborators	
Invest in training and capability building	Recommendation cited in 2 studies of high quality, 4 of medium quality, and 2 of low quality.
Assess organisational infrastructure before funding	Recommendation cited in 2 studies of high quality and 2 of medium quality.
Assess fit with organisational mission	Recommendation cited in 2 studies of high quality and 1 of medium quality.
Planning and communication of costs before implementation	Recommendation cited in 2 studies of medium and high quality.
Ensure partnership with community of implementation	Recommendation cited in 2 studies of medium and high quality.
Assess organisational needs of TSOs	Recommendation cited in 2 studies of medium and high quality.
Provide continuous evaluation and support	Recommendation cited in 2 studies of medium and high quality.

Recommendations for research and practice	
Clearer guidelines to how an intervention can be changed to respond to agency and client needs.	Recommendation cited in 2 studies of high quality and 3 of medium quality.
Better understand how TSOs can appropriately modify/adapt interventions	Recommendation cited in 2 studies of high quality and 3 of medium quality.
Request guidance on adaptation versus fidelity by developers.	Recommendation cited in 2 studies of high quality and 1 of medium quality.
To enable and support CBOs in being able to adapt and tailor interventions.	Recommendation cited in 2 studies of medium and high quality.
Recommendations for practitioners	
Ensure necessary capability to implement and adapt	Recommendation cited in 3 studies of high quality and 3 of medium quality.
Invest in organizational infrastructure	Recommendation cited in 3 studies of medium quality.
Preparation to implement	Recommendation cited in 2 studies of medium and high quality.
Maintain and build expertise	Recommendation cited in 2 studies of medium and high quality.

The main recommendations for *funders* are to invest in technical assistance and capability training for the TSOs they fund, and to assess organizational infrastructure and ability of TSOs to implement EBIs before offering funding. The main recommendation for *research and practice* are to have clearer guidelines on how to adapt and modify EBIs to different populations, and to conduct more research on how EBIs can be modified without compromising effectiveness. The main recommendations for *practitioners* are to ensure capability to implement and adapt EBIs, and to invest in organizational infrastructure. These results will be addressed further in the discussion.

Discussion

Summary of findings

This systematic review identified, quality-appraised and synthesised 31 studies, most of which were conducted in the US. The thematic synthesis identified the most reported barrier for TSOs to implement EBIs to revolve around *resources*, in particular with respect to lack of staff and finance. The most reported category of facilitators related to *organisational culture* and included factors such as alignment between the mission of the TSO and the EBI, and organisational

support for implementing the EBI. The findings were largely representative of the included studies and robust to study quality.

Implications for policy and practice

In interpreting the thematic categories, one may distinguish between *capacity* and *capability*. Capacity can be thought of as 'how much' an organisation can do and include aspects related to resources and infrastructure, whereas capability can be understood as 'how well' an organisation can do and thus relates to issues such as expertise and experience [88, 89]. Importantly, the review demonstrated that TSO practitioners experience central issues related to both capacity and capability. Specifically, the most salient barriers were related to a lack of resources, issues with clients and community (particularly in terms of recruitment and retention issues), and a lack of delivery expertise and experience.

These barriers were also reflected in the recommendations, which highlighted a need for funders to better assess TSOs before providing funding and, when having made the decision to fund an organisation, to provide continuous training and support for that organisation. For practitioners, the main recommendation was to invest in the necessary organisational infrastructure and to ensure that organisations have the technical expertise to implement an EBI. This suggests that funders often do not provide the necessary support for TSOs to implement EBIs, but also that more emphasis should be put on selecting the TSOs that are most capable and motivated to becoming evidence-based providers.

A reoccurring theme for both barriers, facilitators and recommendations was that of *adaptation*, in which support for adaptation of EBIs was mentioned as an important facilitator, and the lack of adaptation guidelines as a central barrier. Many EBIs, such as DEBIs, come in packages with clear guidelines on how they should be implemented, but do not always address the issue of

adaptation and modification [90–92]. Many studies reported that the *one-size-fits-all* form of EBIs often constituted a challenge, particularly regarding being unable to appropriately adapt an intervention to the target population of the TSO. This challenge is, for example, mirrored by the most reported impeding factor being 'recruitment/retention issues'. Further, several studies reported that TSOs did not receive sufficient training and support in learning how to appropriately adapt the interventions. Some of the challenges around adaptation are reflected in the following quote:

> "The question of fidelity is something that they talked about a lot… the boxed [DEBI] interventions are great, but what people really need is more technical assistance about how to effectively adapt these interventions while retaining the theoretical core. They [the agency] needs to build their capacity to understand the internal logic of the M-group piece of the intervention so that they can say 'here is the logic of this activity, and the behavior it is seeking to address…here is our target population for this intervention…how do we change M-groups for this target population while retaining fidelity to the original design?'" [19]

Past research has already emphasised the importance of theorising and outlining the 'core components' of interventions as part of implementation guidelines [93, 94]. However, even in clinical practice, appropriate and sufficient reporting on intervention components remains an issue [37, 95, 96]. Whilst efforts are arguably being made to improve the reporting of interventions (e.g. through the TIDieR checklist [97]), this is a fairly recent development, and many EBIs, such as DEBIs, do not always include sufficient descriptions of intervention components [91]. Also, simply reporting on intervention components is not enough; efforts must also be made to consider whether, and if so, how, different components can be effectively adapted to different contexts and populations [98].

The challenge of EBIs sometimes being too inflexible in their form was also mirrored in the recommendations for research and practice, which proposed to better understand how one can effectively modify interventions and to provide clearer guidelines on how to adapt EBIs to different populations. Importantly, this does not mean to suggest that one should strive for *standardisation* of intervention components (e.g. activities), but rather that more efforts are needed to conceptualise how intervention functions or mechanisms can be tailored to different implementation settings [99].

Ways forward

This systematic review identified the categories reported as barriers and facilitators to be largely distinct, rather than reciprocal which is often the case in this type of review [58, 59, 100]. For example, the most reported barrier, resources, was one of the least reported factors operating as a facilitator. This finding may appear curious, considering that the included studies seem to put heavy emphasis on the lack of resources as a barrier; yet 'more resources' was not often identified as a facilitating factor. Similarly, by far the most reported facilitating category, organisational culture, was only the fourth most reported theme of categories operating as barriers.

One may think about this non-reciprocal relationship between barriers and facilitators through the scope of sufficient and necessary factors. That is, to improve practice, it may be necessary to provide more resources for implementation, but it may not be a sufficient step in isolation. As demonstrated by the analysis of facilitators, demonstrating internal motivation to prioritise the EBI was perceived central for successful implementation. However, regardless of the degree of organisational motivation and support, an organisation must have sufficient resources in place to become an evidence-based provider.

The idea that one must consider the implementation process in its entirety was also reflected in the list of recommendations, which entailed different suggestions for different stakeholders. Past research has consistently reported that TSOs are dependent on the requirements posed by funders and commissioners [101–103], and that their ability to adhere to best practice is largely a function of those criteria. Thus, to improve current practice it is central to ensure that commissioning criteria are relevant and facilitating of the best possible practice.

While an increasing body of academic and policy literature contends that the third sector plays an important role in the future delivery of care and health, there seems to be limited attention to how the role of TSOs can be best facilitated. For example, little is known about whether TSOs require specific support in the delivery of EBIs compared to other types of providers. However, TSOs often report that the commissioning requirements imposed on them favour a New Public Management line of thinking, thus focussing on financial reporting and management [106, 107], which, in isolation, may overlook central aspects related to implementation and effectiveness.

To respond to the challenges identified in this review, it is important to consider perspectives from the full system of stakeholders, including practitioners, commissioners, service-users, and policymakers to discuss how TSOs can be best supported in delivering social and health services. To do so, one may draw on stakeholder inclusive methods such as consensus meetings and Delphi panels [108, 109] to determine what types of support TSOs require to appropriately deliver EBIs. Importantly, this would also allow for identifying the organisational *needs* and *constraints* of TSOs and to engage in dialogue about how the commissioning process may better assess the potential of organisations to deliver EBIs, and what types of support will best enable successful implementation.

Limitations

Most of the included studies were conducted in North America, and the findings of this review should be interpreted in the light of that context. Similarly, the review aggregated factors from a mixed sample of organisations within health and social care, and thus mainly applies to organisations of those types. Further, the review suffers from the fundamental limitation that different factors may affect the implementation of different organisations or EBIs differently. However, the studies which did investigate barriers and facilitators across different EBIs tended to identify factors that revolved around similar themes [30, 60, 76].

The synthesis was done by aggregating identified factors, and thus followed a sort of 'vote-counting'. However, by analysing and categorising the identified factors thematically, the findings should be more robust to factors revolving around the same underlying concept [110]. Also, by mapping the contribution to the categories of the included studies, the review demonstrated how representative the constructed themes were of the included research. Further, all identified studies were quality appraised, and the reliability of the identified factors were assessed according to both the consistency with which they were reported and the quality of studies from which the factors originated.

Further research

This review focused on the perspectives of practitioners following the assumption that their experiences were closest to the implementation process, but future research may consider a larger group of stakeholders, such as commissioners and policymakers. The review demonstrated that many implementation issues were grounded in the lack of adaptation expertise, driven in turn by the lack of support and expertise, and unclear/insufficient guidelines on how to adapt EBIs to different populations. To address this challenge, future research should investigate how to provide clearer guidelines for TSOs to adapt EBIs and what types of support the adaptation

process requires. However, such investigations mandate careful thought on and clear outlining of what types of modifications can be made without compromising core intervention components.

This review focused on factors influencing implementation and did not discuss how these factors might correspond to other aspects, such as *evaluation* [111]. Becoming an evidence-based provider does not just entail implementing services developed using the best available evidence and stakeholder preferences [112], but also mandate the use of *continuous* monitoring and evaluation [113, 114]. However, if an intervention is developed and tested according to best practice, is it then necessary to allocate further resources to evaluation? One might argue that requiring continuous evaluation on interventions which have already proved efficacious takes away valuable resources from *service delivery*. To approach this discussion, one might consider the example of mass de-worming programmes, which involve medicating children against certain soil-transmitted infections in endemic areas [115]. These programmes have been – and still are – heavily implemented following an assumption of effectiveness based on single studies demonstrating promising results [116]. Yet, the utility of these interventions were recently questioned by two high quality systematic reviews which both demonstrated null effects of the de-worming intervention [117, 118]. Such findings emphasise the importance of continuous evaluation, especially for interventions that have only proved efficacious in single studies. However, to engage in recommendations and guidelines on how and when implementation might influence effectiveness, more research is warranted to better theorise how implementation may moderate intervention effects, and how commissioning criteria may consider such moderations.

Conclusion

Whilst it is often assumed that good outcomes follow when implementing interventions that have been developed and tested according to the principles of EBP, little attention has been paid

to how EBIs are best transported to different service settings [119]. Generally, past research has arguably taken the aspect of implementation for granted, even in clinical settings [36, 120–122]. However, particularly little is known about the implementation of third sector service deliveries, which is a central blind spot considering the increasing role of TSOs in the delivery of social and health services.

This review constitutes the first systematic attempt to aggregate and analyse the factors that influence the implementation process of TSOs. In so doing, this review addresses the often-implicit assumption that interventions and programmes supported by rigorous evidence can be effectively implemented without considering the implementation ability of service providers. The findings illustrate that TSOs face many capacity and capability issues, which are primarily driven by a lack of support and expertise. Going forward, it is central to involve and engage with stakeholders to discuss how the commissioning process may better identify capable TSOs and offer the necessary implementation support.

List of abbreviations: Evidence-based intervention: EBI, evidence-based practice: EBP, third sector organisation: TSO, community based organisations: CBO, United States: US

Acknowledgements: We are grateful to Kathryn Oliver and Jane Barlow for providing helpful comments to the review protocol.

Bibliography

1. Despard MR. Challenges in Implementing Evidence-Based Practices and Programs in Nonprofit Human Service Organizations. J Evidence-Informed Soc Work. 2016;13:505–22. doi:10.1080/23761407.2015.1086719.
2. Dickinson H, Allen K, Alcock P, Macmillan R, Glasby J. The Role of the Third Sector in Delivering Social Care. NIHR School for Social Care Research. 2012. doi:978-0-85328-443-7.
3. Macmillan R. The third sector delivering public services: an evidence review. TSRC Working Paper 20 (Birmingham, University of Birmingham). 2010. https://www.birmingham.ac.uk/generic/tsrc/research/service-delivery/wp-20-public-services-evidence-review.aspx. Accessed 7 Apr 2016.

4. Wilson M, Lavis J, Guta A. Community-based organizations in the health sector: A scoping review. Heal Res Policy Syst. 2012;10:36. doi:10.1186/1478-4505-10-36.
5. Hardwick R, Anderson R, Cooper C. How do third sector organisations use research and other knowledge? A systematic scoping review. Implement Sci. 2015;10:84. doi:10.1186/s13012-015-0265-6.
6. Chapman T, Brown J, Crow R. Entering a brave new world? An assessment of third sector readiness to tender for the delivery of public services in the United Kingdom. Policy Stud. 2008;29:1–17. doi:10.1080/01442870701847998.
7. Alcock P. A strategic unity: defining the third sector in the UK. Volunt Sect Rev. 2010;1:5–24.
8. Defourny J, Kirsten Grønbjerg B, Lucas Meijs B, Marthe Nyssens B, Naoto Yamauchi B, Grønbjerg K, et al. Voluntas Symposium: Comments on Salamon and Sokolowski's Re-conceptualization of the Third Sector. VOLUNTAS. 2016;27:1546–61. doi:10.1007/s11266-016-9743-y.
9. Anheier HK. Nonprofit Organizations: Theory, management, policy. Routledge; 2005. http://books.google.com/books?id=s_uzg-KoVyIC&pgis=1.
10. Morris S. Defining the non-profit sector: Some lessons from history. 2000. http://eprints.lse.ac.uk/29032/1/cswp3.pdf. Accessed 29 Mar 2017.
11. Rees J, Mullins D, editors. The third sector delivering public services : developments, innovations and challenges. Policy Press; 2016.
12. Barron DN, West E. The quasi-market for adult residential care in the UK: Do for-profit, not-for-profit or public sector residential care and nursing homes provide better quality care? Soc Sci Med. 2017;179:137–46. doi:10.1016/j.socscimed.2017.02.037.
13. Moore MH. Managing for Value: Organizational Strategy in for-Profit, Nonprofit, and Governmental Organizations. Nonprofit Volunt Sect Q. 2000;29:183–204. doi:10.1177/089976400773746391.
14. Dart R. Being "Business-Like" in a Nonprofit Organization: A Grounded and Inductive Typology. Nonprofit Volunt Sect Q. 2004;33:290–310.
15. Hogg E, Baines S. Changing Responsibilities and Roles of the Voluntary and Community Sector in the Welfare Mix: A Review. Soc Policy Soc. 2011;10:341–52. doi:10.1017/S1474746411000078.
16. Kelly J. Reforming public services in the UK: Bringing in the third sector. Public Adm. 2007;85:1003–22. doi:10.1111/j.1467-9299.2007.00680.x.
17. Cabinet Office. Open Public Services White Paper. 2011. http://www.cabinetoffice.gov.uk/sites/default/files/resources/open-public-services-white-paper.pdf%5Cnhttp://www.openpublicservices.cabinetoffice.gov.uk/.
18. Office for Civil Society. Building a Stronger Civil Society. 2010.
19. Kegeles SM, Rebchook G, Tebbetts S, Arnold E. Facilitators and barriers to effective scale-up of an evidence-based multilevel HIV prevention intervention. Implement Sci. 2015;10:50. doi:10.1186/s13012-015-0216-2.
20. Chapman T, Bell V, Robinson F. Measuring impact: easy to say, hard to do: a think-piece to stimulate sector debate from the Third Sector Trends Study. 2011;:17.
21. Gibbs L, Gambrill E. Evidence-Based Practice: Counterarguments to Objections. Res Soc Work Pract. 2002;12:452–76. doi:10.1177/1049731502012003007.
22. Sackett DL, Rosenberg WMC, Gray J a M, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. Br Med J. 1996;312:71–2.

23. Fraser MW, Galinsky MJ. Steps in Intervention Research: Designing and Developing Social Programs. Res Soc Work Pract. 2010;20:459–66.
24. Magill M. The Future of Evidence in Evidence-based Practice: Who Will Answer the Call for Clinical Relevance? J Soc Work. 2006;6:101–15. doi:10.1177/1468017306066737.
25. What Works? 2014. www.nice.org.uk. Accessed 29 May 2017.
26. Project Oracle. http://project-oracle.com/.
27. Blueprints for Healthy Child Development. http://www.blueprintsprograms.com/.
28. Effective Interventions. https://effectiveinterventions.cdc.gov/. Accessed 17 Nov 2017.
29. Margaret Dolcini M, Gandelman AA, Vogan SA, Kong C, Leak TN, King AJ, et al. Translating HIV interventions into practice: Community-based organizations' experiences with the diffusion of effective behavioral interventions (DEBIs). Soc Sci Med. 2010;71:1839–46.
30. Amodeo M, Lundgren L, Cohen A, Rose D, Chassler D, Beltrame C, et al. Barriers to implementing evidence-based practices in addiction treatment programs: Comparing staff reports on Motivational Interviewing, Adolescent Community Reinforcement Approach, Assertive Community Treatment, and Cognitive-behavioral Therapy. Eval Program Plann. 2011;34:382–9. doi:10.1016/j.evalprogplan.2011.02.005.
31. Ubbink DT, Guyatt GH, Vermeulen H. Framework of policy recommendations for implementation of evidence-based practice: a systematic scoping review. BMJ Open. 2013;3:e001881-. doi:10.1136/bmjopen-2012-001881.
32. Gray M, Joy E, Plath D, Webb SA. What supports and impedes evidence-based practice implementation? A survey of Australian social workers. Br J Soc Work. 2013;45:667–84. doi:10.1093/bjsw/bct123.
33. Montgomery P, Underhill K, Gardner F, Operario D, Mayo-Wilson E. The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. J Clin Epidemiol. 2013;66:874–82. doi:10.1016/j.jclinepi.2013.03.006.
34. Durlak JA, DuPre EP. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. Am J Community Psychol. 2008;41:327–50.
35. Hasson H. Systematic evaluation of implementation fidelity of complex interventions in health and social care. Implement Sci. 2010;5:67. doi:10.1186/1748-5908-5-67.
36. Durlak JA, Wells AM. Primary prevention mental health programs for children and adolescents: a meta-analytic review. Am J Community Psychol. 1997;25:115–52.
37. Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? BMJ Br Med J. 2008;336:1472–4. doi:10.1136/bmj.39590.732037.47.
38. Mihalic SF. The Importance of Implementation Fidelity. Emot Behav Disord youth. 2004;4:83–105. doi:10.1017/CBO9781107415324.004.
39. Basch CE, Sliepcevich EM, Gold RS, Duncan DF, Kolbe LJ. Avoiding type III errors in health education program evaluations: a case study. Health Educ Q. 1985;12:315–31. http://www.ncbi.nlm.nih.gov/pubmed/4077544. Accessed 17 Mar 2017.
40. Dobson D, Cook TJ. Avoiding type III error in program evaluation. Eval Program Plann. 1980;3:269–76. doi:10.1016/0149-7189(80)90042-7.
41. Hasson H. Systematic evaluation of implementation fidelity of complex interventions in health and social care. 2010. doi:10.1186/1748-5908-5-67.
42. McCord J. Cures That Harm: Unanticipated Outcomes of Crime Prevention Programs. Ann

Am Acad Pol Soc Sci. 2003;587:16–30.
43. Dishion TJ, Mccord J, Poulin F. When Interventions Harm: Peer Groups and Problem Behavior. Am Psychol . 1999;54:755–64. doi:10.1037/0003-066X.54.9.755.
44. Petrosino A, Turpin-petrosino C, Hollis-peel ME, Lavenberg JG. 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency. Cochrane Database Syst Rev. 2013;:4–6.
45. NCVO. NCVO UK Civil Society Almanac. 2017. https://data.ncvo.org.uk/. Accessed 5 Jan 2017.
46. Lundgren L, Krull I, Zerden L de S, Mccarty D. Community-based addiction treatment staff attitudes about the usefulness of evidence-based addiction treatment and CBO organizational linkages to research institutions. Eval Program Plann. 2011;34:356–65. doi:10.1016/j.evalprogplan.2011.02.002.
47. Ramanadhan S, Crisostomo J, Alexander-Molloy J, Gandelman E, Grullon M, Lora V, et al. Perceptions of evidence-based programs among community-based organizations tackling health disparities: a qualitative study. Health Educ Res. 2012;27:717–28. doi:10.1093/her/cyr088.
48. Owczarzak J, Dickson-Gomez J. Provider perspectives on evidence-based HIV prevention interventions: barriers and facilitators to implementation. AIDS Patient Care STDS. 2011;25:171–9.
49. Martínez G, Sardiñas LM, Acosta-Perez E, Medina L, Rivera M, Pattatucci A. Capacity needs in community-based organizations for enhancing translational research in Puerto Rico. Prog Community Health Partnersh. 2014;8:53–60. https://search.proquest.com/docview/1529846240?accountid=13042 http://oxfordsfx.hosted.exlibrisgroup.com/oxford?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&genre=article&sid=ProQ:ProQ%3Amedlinesh ell&atitle=Capacity+needs+in+community-base.
50. Lundgren L, Amodeo M, Cohen A, Chassler D, Horowitz A, Schoenwald S, et al. Modifications of evidence-based practices in community-based addiction treatment organizations: A qualitative research study. Addict Behav. 2011;36:630–5. doi:10.1016/j.addbeh.2011.01.003.
51. Kimber M, Barwick M, Fearing G. Becoming an Evidence-Based Service Provider: Staff Perceptions and Experiences of Organizational Change. J Behav Health Serv Res. 2012;39:314–32. doi:10.1007/s11414-012-9276-0.
52. Mullen E, Shlonsky A. From concept to implementation: Challenges facing evidence-based social work. Evid Policy. 2005;1:61–84. http://www.ingentaconnect.com/content/tpp/ep/2005/00000001/00000001/art00004.
53. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. Int J Surg. 2010;8:336–41.
54. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015;349. http://www.bmj.com/content/349/bmj.g7647. Accessed 16 Aug 2017.
55. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev. 2016;5:210. doi:10.1186/s13643-016-0384-4.
56. Thomas J, Harden A, Newman M. Synthesis: Combining results systematically and appropriately. 1st edition. London: Sage; 2012.

57. Tricco AC, Cardoso R, Thomas SM, Motiwala S, Sullivan S, Kealey MR, et al. Barriers and facilitators to uptake of systematic reviews by policy makers and health care managers: a scoping review. Implement Sci. 2016;11:4. doi:10.1186/s13012-016-0370-1.
58. Gravel K, Légaré F, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: a systematic review of health professionals' perceptions. Implement Sci. 2006;1:16. doi:10.1186/1748-5908-1-16.
59. Oliver K, Innvar S, Lorenc T, Woodman J, Thomas J. A systematic review of barriers to and facilitators of the use of evidence by policymakers.(Research article). BMC Health Serv Res. 2014;14:2.
60. Demby H, Gregory A, Broussard M, Dickherber J, Atkins S, Jenner LW. Implementation lessons: The importance of assessing organizational "fit" and external factors when implementing evidence-based teen pregnancy prevention programs. J Adolesc Heal. 2014;54 3 SUPPL.:S37–44. doi:10.1016/j.jadohealth.2013.12.022.
61. Rees R, Oliver K, Woodman J, Thomas J. Children's views about obesity, body size, shape and weight: a systematic review. 2009. http://eppi.ioe.ac.uk/cms/Portals/0/Obesity Views Children R2009Rees.pdf?ver=2010-12-22-121209-040. Accessed 3 Apr 2017.
62. Verboom B, Montgomery P, Bennett S. What factors affect evidence-informed policymaking in public health? Protocol for a systematic review of qualitative evidence using thematic synthesis. Syst Rev. 2016;5. doi:10.1186/s13643-016-0240-6.
63. Maharaj R. Organizational Culture, Absorptive Capacity, and the Change Process: Influences on the Fidelity of Implementation of Integrated Dual Disorder Treatment in Community-Based Mental Health Organizations A. 2010.
64. Pemberton GC. How are implementation and adaptation of evidence-based interventions applied in community practice settings? lessons from the modelo de intervencion psicomedica. 2012. http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2013-99060-103&site=ehost-live.
65. Belza B, Petrescu-Prahova M, Kohn M, Miyawaki CE, Farren L, Kline G, et al. Adoption of Evidence-Based Health Promotion Programs: Perspectives of Early Adopters of Enhance®Fitness in YMCA-Affiliated Sites. Front Public Heal. 2015;2 April:1–12. doi:10.3389/fpubh.2014.00164.
66. Dippolito M, Lundgren L, Amodeo M, Beltrame C, Lim L, Chassler D. Addiction treatment staff perceptions of training as a facilitator or barrier to implementing evidence-based practices: A national qualitative research study. Subst Abus. 2015;36:42–50.
67. Flores R, Naranjo C, Hein A. Use of evidence in the implementation of social programs: A qualitative study from Chile. J Evidence-Informed Soc Work. 2016;13:562–75.
68. Gandelman A, Dolcini MM. The influence of social determinants on evidence-based behavioral interventions-considerations for implementation in community settings. Transl Behav Med. 2012;2:137–48.
69. Kimber M, Barwick M, Fearing G. Becoming an evidence-based service provider: staff perceptions and experiences of organizational change. J Behav Heal Serv Res. 2012;39 July:314–32. doi:10.1007/s11414-012-9276-0.
70. Lattimore D, Griffin SF, Wilcox S, Rheaume C, Dowdy DM, Leviton LC, et al. Understanding the challenges encountered and adaptations made by community organizations in translation of evidence-based behavior change physical activity interventions: A qualitative study. Am J Heal Promot. 2010;24:427–34.

71. Owczarzak J. Evidence-based HIV prevention in community settings: provider perspectives on evidence and effectiveness. Crit Public Health. 2012;22:73–84.
72. Payan DD, Sloane DC, Illum J, Vargas RB, Lee D, Galloway-Gilliam L, et al. Catalyzing Implementation of Evidence-Based Interventions in Safety Net Settings: A Clinical-Community Partnership in South Los Angeles. Health Promot Pract. 2017;18:586–97.
73. Petrescu-Prahova M, Belza B, Kohn M, Miyawaki C. Implementation and maintenance of a community-based older adult physical activity program. Gerontologist. 2016;56:677–86.
74. Pinto BM, Waldemore M, Rosen R. A Community-Based Partnership to Promote Exercise Among Cancer Survivors: Lessons Learned. Int J Behav Med. 2015;22:328–35.
75. Veniegas RC, Kao UH, Rosales R. Adapting HIV prevention evidence-based interventions in practice settings: an interview study. Implement Sci. 2009;4:76. doi:10.1186/1748-5908-4-76.
76. Collins CB, Hearn KD, Whittier DN, Freeman A, Stallworth JD, Phields M. Implementing packaged HIV-prevention interventions for HIV-positive individuals: considerations for clinic-based and community-based interventions. Public Health Rep. 2010;125 Suppl:55–63. doi:10.2307/41434871.
77. Murray M, Culver T, Farmer B, Jackson LA, Rixon B. From Theory to Practice: One Agency's Experience with Implementing an Evidence-Based Model Maureen. J Child Fam Stud. 2014;:844–853.
78. Feutz K, Andresen P. Cuidate: Implementation of a Culturally Based Sexual Risk Reduction Program for Hispanic Adolescents. Hisp Heal Care Int. 2013;11:187–93. doi:10.1891/1540-4153.11.4.187.
79. Honeycutt S, Carvalho M, Glanz K, Daniel SD, Kegler MC. Research to Reality. J Public Heal Manag Pract. 2012;18:431–9. doi:10.1097/PHH.0b013e31822d4c69.
80. House LD, Tevendale HD, Martinez-Garcia G. Implementing Evidence-Based Teen Pregnancy-Prevention Interventions in a Community-Wide Initiative: Building Capacity and Reaching Youth. J Adolesc Heal. 2017;60:S18–23. doi:10.1016/j.jadohealth.2016.08.013.
81. Lundgren L, Chassler D, Amodeo M, D'Ippolito M, Sullivan L. Barriers to implementation of evidence-based addiction treatment: a national study. J Subst Abuse Treat. 2012;42:231–8. doi:10.1016/j.jsat.2011.08.003.
82. Thomas R, Zimmer-Gembeck MJ, Chaffin M. Practitioners' views and use of evidence-based treatment: Positive attitudes but missed opportunities in children's services. Adm Policy Ment Heal Ment Heal Serv Res. 2014;41:368–78.
83. Hunter SB, Han B, Slaughter ME, Godley SH, Garner BR. Predicting evidence-based treatment sustainment: results from a longitudinal study of the Adolescent-Community Reinforcement Approach. Implement Sci. 2017;12:75. doi:10.1186/s13012-017-0606-8.
84. Allicock M, Johnson L-S, Leone L, Carr C, Walsh J, Ni A, et al. Promoting Fruit and Vegetable Consumption Among Members of Black Churches, Michigan and North Carolina, 2008–2010. Prev Chronic Dis. 2013;10:120161. doi:10.5888/pcd10.120161.
85. Vanderpool R, Gainor S, Conn M, Spencer C, Allen A, Kennedy S. Rural Remote Health. Rural Remote Heal. 2011;193:118–25.
86. Kegeles SM, Rebchook G, Pollack L, Huebner D, Tebbetts S, Hamiga J, et al. An Intervention to Help Community-Based Organizations Implement an Evidence-Based HIV Prevention Intervention: The Mpowerment Project Technology Exchange System. Am J Community Psychol. 2012;49:182–98.
87. Lundgren L, Amodeo M, Cohen A, Chassler D, Horowitz A, Schoenwald S, et al.

Modifications of evidence-based practices in community-based addiction treatment organizations: A qualitative research study. Addict Behav. 2011;36:630–5. doi:10.1016/j.addbeh.2011.01.003.
88. Macmillan R, Ellis A, With P, Kara H, Dayson C, Sanderson E, et al. Building Capabilities in the Voluntary Sector: What the evidence tells us. TSRC Research Report 125 (Birmingham, University of Birmingham). 2014. https://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/reports/research-report-125-building-capabilities.pdf. Accessed 9 Dec 2017.
89. Building capacity: Research - Big Lottery Fund. https://www.biglotteryfund.org.uk/research/making-the-most-of-funding/building-capacity. Accessed 14 Dec 2017.
90. Collins C, Harshbarger C, Sawyer R, Hamdallah M. The diffusion of effective behavioral interventions project: development, implementation, and lessons learned. AIDS Educ Prev. 2006;18 4 Suppl A:5–20. doi:10.1521/aeap.2006.18.supp.5.
91. Collins CB, Sapiano TN. Lessons Learned From Dissemination of Evidence-Based Interventions for HIV Prevention. Am J Prev Med. 2016;51:S140–7. doi:10.1016/j.amepre.2016.05.017.
92. Peterson AS, Randall LM. Utilizing multilevel partnerships to build the capacity of community-based organizations to implement effective HIV prevention interventions in Michigan. AIDS Educ Prev. 2006;18 SUPPL. A:83–95.
93. Michie S, Fixsen D, Grimshaw JM, Eccles MP. Specifying and reporting complex behaviour change interventions: the need for a scientific method. Implement Sci. 2009;4:40. doi:10.1186/1748-5908-4-40.
94. Gearing RE, El-Bassel N, Ghesquiere A, Baldwin S, Gillies J, Ngeow E. Major ingredients of fidelity: A review and scientific guide to improving quality of intervention research implementation. Clin Psychol Rev. 2011;31:79–88. doi:10.1016/j.cpr.2010.09.007.
95. Weersing VR, Rozenman M, Gonzalez A. Core Components of Therapy in Youth. Behav Modif. 2009;33:24–47. doi:10.1177/0145445508322629.
96. Maggin DM, Johnson AH. The Reporting of Core Program Components: An Overlooked Barrier for Moving Research Into Practice. Prev Sch Fail Altern Educ Child Youth. 2015;59:73–82. doi:10.1080/1045988X.2013.837812.
97. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348 mar07_3:g1687. doi:10.1136/bmj.g1687.
98. Hawe A. T., P Shiell R. Complex interventions: how "out of control" can a randomised controlled trial be? Bmj. 2004;328 June:1561–3.
99. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. 2011. doi:10.1186/1748-5908-6-42.
100. Légaré F, Ratté S, Gravel K, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: Update of a systematic review of health professionals' perceptions. Patient Educ Couns. 2008;73:526–35.
101. Mitchell GE, Berlan D. Evaluation and Evaluative Rigor in the Nonprofit Sector. Nonprofit Manag Leadersh. 2016;27:237–50.
102. Harlock J. Impact measurement practice in the UK third sector: A review of emerging evidence. TSRC Working Paper 106 (Birmingham, University of Birmingham). 2013.

http://epapers.bham.ac.uk/1800/1/WP106_Impact_measurement_practice_in_the_UK_third_sector_-_Harlock%2C_July_2013.pdf. Accessed 7 Apr 2016.
103. Arvidson M. Evidence and transparency in the open public services reform: perspectives for the third sector. TSRC Working Paper 117 (Birmingham, University of Birmingham). 2014. https://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-117.pdf. Accessed 6 Aug 2016.
104. Rutherford AC. Rising Wages in the Expanding U.K. Nonprofit Sector From 1997 to 2007. Nonprofit Volunt Sect Q. 2015;44:123–45. doi:10.1177/0899764013509891.
105. Cabinet Office. The future role of the third sector in social and economic regeneration: final report. 2007. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/228900/7189.pdf. Accessed 24 Apr 2018.
106. Carmel E, Harlock JE. Instituting the "third sector" as a governable terrain: Partnership, procurement and performance in the UK. Policy Polit. 2008;36:155–71.
107. Maier F, Meyer M, Steinbereithner M. Nonprofit Organizations Becoming Business-Like: A Systematic Review. Nonprofit Volunt Sect Q. 2016;45:64–86. doi:10.1177/0899764014561796.
108. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. PLoS Med. 2010;7:e1000217. doi:10.1371/journal.pmed.1000217.
109. Hsu C, Sandford B. The delphi technique: making sense of consensus. Pract Assessment, Res Eval. 2007;12:1–8.
110. Thomas J, Harden A, Thomas,James, Harden,Angela. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol. 2008;8:45. doi:10.1186/1471-2288-8-45.
111. Bach-Mortensen AM, Montgomery P. What are the barriers and facilitators for third sector organisations (non-profits) to evaluate their services? A systematic review. Syst Rev. 2018;7:13. doi:10.1186/s13643-018-0681-1.
112. Sackett D. Evidence-based medicine : how to practice and teach EBM. Churchill Livingstone; 2000. doi:10.1016/S1031-170X(97)80036-0.
113. Chalmers I. If evidence-informed policy works in practice, does it matter if it doesn't work in theory? Evid Policy. 2005;1:227–42.
114. Chalmers I. Trying to do more Good than Harm in Policy and Practice: The Role of Rigorous, Transparent, Up-to-Date Evaluations. Ann Am Acad Pol Soc Sci. 2003;589:22–40. doi:10.1177/0002716203254762.
115. WHO | Deworming in children. WHO. 2018. http://www.who.int/elena/titles/deworming/en/. Accessed 22 Apr 2018.
116. Miguel E, Kremer M. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. Econometrica. 2004;72:159–217. doi:10.1111/j.1468-0262.2004.00481.x.
117. Welch VA, Ghogomu E, Hossain A, Awasthi S, Qar Z, Bhutta A, et al. Mass deworming to improve developmental health and wellbeing of children in low-income and middle-income countries: a systematic review and network meta-analysis. 2017;5. doi:10.1016/S2214-109X(16)30242-X.
118. Taylor-Robinson DC, Maayan N, Soares-Weiser K, Donegan S, Garner P. Deworming drugs for soil-transmitted intestinal worms in children: effects on nutritional indicators, haemoglobin, and school performance. In: Cochrane Database of Systematic Reviews. Taylor-

Robinson DC, editor. Chichester, UK: John Wiley & Sons, Ltd; 2015. doi:10.1002/14651858.CD000371.pub6.
119. Mchugh RK, Barlow DH, Kathryn R. The Dissemination and Implementation of Evidence-Based Psychological Treatments. 2010;65:73–84.
120. Proctor EK, Landsverk J, Aarons G, Chambers D, Glisson C, Mittman B. Implementation research in mental health services: An emerging science with conceptual, methodological, and training challenges. Adm Policy Ment Heal Ment Heal Serv Res. 2009;36:24–34. doi:10.1007/s10488-008-0197-4.
121. Groene O, Botje D, Suñol R, Lopez MA, Wagner C. A systematic review of instruments that assess the implementation of hospital quality management systems. Int J Qual Heal Care. 2013;25:525–41.
122. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Heal Ment Heal Serv Res. 2011;38:65–76. doi:10.1007/s10488-010-0319-7.

Chapter 5: Does sector matter for the quality of care services? A secondary analysis of social care services regulated by the Care Inspectorate in Scotland[5]

Anders Malthe Bach-Mortensen, MSc[a]

Paul Montgomery, DPhil, MSc, DipSW[b]

[a] University of Oxford, Department of Social Policy and Intervention; Barnett House, 32 Wellington Square; Oxford; UK

[b] University of Birmingham, Department of Social Policy, Sociology and Criminology, Birmingham, B15 2TT, UK.

Author contributions: The research design and objectives were developed by ABM and PM. ABM designed and conducted the statistical analysis and drafted the manuscript with input from PM. All authors read and approved the final manuscript.

[5] This paper has been published in BMJ Open.

Abstract

Objectives: Social services are increasingly commissioned to third and for-profit sector providers, but little is known about whether and how these changes influence quality indicators. We assessed quality-related outcomes across for-profit, public and third sector organisations delivering social care services.

Design: A secondary analysis was conducted on publically available data collected by the independent regulator of social care organisations in Scotland. All outcomes are reported as predicted probabilities derived from multivariate logistic regression coefficients. Generalised ordered logit models are utilised for the quality domains and the risk assessment score, and logistic regression for whether complaints or requirements were issued to organisations.

Setting: Organisations inspected by the Care Inspectorate in Scotland.

Population: 13,310 social care organisations (e.g. nursing homes and day care organisations).

Primary outcomes: The quality and risk domains collected by the Care Inspectorate and complaints and requirements issued to organisations within the last three years.

Results: Controlling for multiple factors, we find that public and third sector providers performed consistently and statistically significantly better than for-profit organisations on most outcomes. For example, for-profit services were the most likely to be rated as high and medium risk (6.9% and 13.2%, respectively), and the least likely to be classified as low risk (79.9%). Public providers had the highest probability of being categorised as low risk (91.1%), and the lowest probability of having their services classified as medium (6.9%) and high risk (2%), followed by third sector providers (86%, 8.5%, and 4.5%, respectively). Public providers performed better than third sector providers in some outcomes, but differences were relatively low and inconsistent.

Conclusion: Public and third sector providers were rated considerably higher than their for-profit counterparts on most observed outcomes. Regulators might utilise this information to consider how social care providers across sector are incentivised to manage their resources.

Acronyms: Third sector organisation: TSO, Public sector organisation: PSO, For-profit organisation: FPO, Care Inspectorate: CI, Care Quality Commission: CQC

Introduction

Health and social care services are often delivered in quasi-markets, and an increasing proportion of these services are implemented by for-profit and third sector organisations (TSOs) [1–4]. Quasi-markets are characterised by free bidding from organisations across different sectors [5], and can be defined as "…markets where the provision of a service is undertaken by competitive providers as in pure markets, but where the purchasers of the service are financed from resources provided by the state instead of from their own private resources." [6] Thus quasi-markets primarily include services that are free or subsidised to service-users (e.g. healthcare or nursing homes), but where providers from all sectors (i.e. public, for-profit and third) compete for service contracts.

It has been argued that quasi-markets can lead to more efficient and equitable service deliveries [5, 7], but there is little conclusive evidence to support that proposition. In theory, quasi-markets may produce favourable outcomes, in that the bidding system can be thought to force organisations to respond to the needs of service-users and allow the best and most innovative providers to attain the contract. However, the utility of this type of market strongly relies on the commissioning structures in place to allocate funding, which is known to have been flawed in the past [8–10]. Further, it has long been understood that mechanisms of action considering behaviour and the design of social policy requires more than simply an analysis of monetary incentives [11].

It has long been assumed that providers from different sectors are subject to different incentives and organisational characteristics [12–16]. For example, for-profit organisations are often perceived to be legitimised by the pursuit of profit maximisation, suggesting that they may engage in opportunistic behaviour at the potential expense of the service-user without regulation. Conversely, third and public sector organisations are often considered to be legitimised by their

ability to achieve *social value,* suggesting that their incentive to deliver services are predominantly motivated by pro-social pursuits, thus making them more 'trustworthy' providers of public services [17]. These theoretical differences across sector are summarised in Table 1.

Table 1: Theoretical distinction of sector

	For-profit sector	Public sector	Third sector
Definition	Organisations that are privately owned and administered according to the pursuit of profits.	Organisations that are publically owned and administered.	Organisations that are formally structured, privately owned, non-profit distributing, self-governing, and benefiting from voluntary activities [18].
Primary activities	Conventionally private goods, but increasingly also public goods	Public goods	Public goods
Main type of income	Profits from sales of products and/or services	Taxation	Charitable contributions and/or commissioning
Purpose/motivation	Maximise and sustain profits	Achieve social mission (public sector motivation/social value)	Achieve social mission and accountability (social value)

Importantly, there is no general consensus on what these theoretical differences between sectors mean in practice and whether, and if so, how, they may relate to quality of service.

To explore sector differences in the context of health services, a relatively recent overview of systematic reviews investigated how ownership influences the performance of healthcare providers [19]. The overview included 15 systematic reviews that were conducted in various contexts (e.g. low and middle income countries (LMICs) and high income countries) and with different types of providers (e.g. hospitals and nursing homes). The paper found mixed results with non-profit providers performing better than for-profit providers on certain indicators (e.g. mortality and price), but not on others (e.g. healthcare related outcomes). Further, the results should be interpreted with the caveat that most of the included systematic reviews had significant methodological flaws and only one of the reviews was considered 'reliable'. As a result, the overview concludes that more high quality primary research is needed, especially in exploring variation across both public, for-profit and third sector organisations, rather than just for-profit and non-profit providers [19].

However, little research has been conducted on sector difference in social care settings, even though more social services are undertaken by for-profit and third sector providers [3, 20, 21], and that sector variation in service quality thus have important implications for practice. To explore the merit of existing theories and to inform future policy on the contracting of social services, it is central to test whether the typology of sector accounts for variation across factors that matter to practice, such as quality and performance.

Objectives

It remains relatively inconclusive whether typology of sector is based on research conventions rather than empirical evidence [22]. The focus of this paper will be to empirically test whether and if so how 'sector' accounts for variation in quality of social care organisations using data from the Care Inspectorate in Scotland. This focus can be summarised by the following research question:

> Does the quality of social care service vary across for-profit, public and third sector organisations as assessed by the quality and risk domains of the Care Inspectorate in Scotland?

Data: The Care Inspectorate

The Care Inspectorate (CI) serves as the regulator of care services in Scotland and currently (January 2018) includes data on 13,310 organisations, which is available for download on the official website and updated monthly [23]. Regulated organisations work within a range of different services, with the biggest categories being child minding, children day-care services, support services, care homes for the elderly, and housing support (see tables A1-A2 in appendix C1 for description of included services).

To fulfil its regulating role, the Care Inspectorate regularly inspects care facilities, which entails assessing the quality and risk of services. The grading of quality is done according to the National Care Standards [24], which considers the following principles: 'Dignity and respect'; 'Compassion'; 'Be inclusive', 'Responsive care and support', and; 'Wellbeing'. The ratings are graded from 1 to 6, in which 1 is unsatisfactory and 6 is excellent. Further, the CI distributes a risk score to care providers assessing whether a service can be said to be of 'low', 'medium' or 'high' risk to service users, which subsequently influences how regularly an organisation should be inspected and the quality domains to which the service should be subjected [25]. The 4 quality domains and the risk assessment are described in Table 2.

Table 2: Quality domains employed by the Care Inspectorate.

Quality and risk domains	Elaboration
Quality of care and support	How well does the service meet the needs of the service-users? This includes looking at the outcomes and materials of the provided services and interviewing service-users.
Quality of Management and Leadership	How is the service managed and does the strategy for development consider the needs of the service-users? This includes interviewing staff and service-users.
Quality of Environment	What is the quality of the setting in which a service is delivered? This includes investigating if the site is clean and properly maintained.
Quality of Staffing	What is the quality of staff in terms of e.g. qualifications and training? This includes interviewing staff and service-users.
Risk of services	A joint assessment of the nature of the services (some services are inherently more 'risky' than others), if complaints have been issued towards a service, and whether any requirements were posed after the last inspection

The inspections are done by trained staff following a standardised procedure, which includes visiting the facility in question unannounced, conducting interviews with staff and service-users, as well as physically examining the facilities [26, 27]. Inspected organisations must also fill out a self-assessment form, which is reviewed in combination with the organisation's history of posed requirements and complaints as part of the grading process. Special guidelines apply to certain care services, such as organisations that work with children and young people [28].

Methods

This paper is reported following best practice guidelines [29, 30]. All analyses were conducted in Stata 14.

The quality and risk domains collected by the CI are ordinal in nature and thus lend themselves to ordered logistic regression. However, this method rests on the *proportional odds* or *parallel lines assumption*, which requires the effect of the odds to be constant across all logits [31]. Substantively this would mean that the effect of sector on quality would be constant across all levels of quality.

We tested this assumption across sector on the quality domains and risk assessment using the 'brant' command [32], which demonstrated the parallel lines assumption to be violated in three quality domains (except for quality of environment) and the risk assessment, thus suggesting that the effect of sector does vary across different levels of quality and risk. To utilise the ordering of the data, while also allowing for the proportional odds assumption to be relaxed, we employ generalised ordered logistic regression, which is less restrictive than ordered logistic regression and more suitable for non-linear associations [31]. Specifically, we will use the gologit2 command written and regularly updated by Richard Williams [33].

Logistic regression is known to entail several limitations, especially when it comes to accurately presenting and interpreting odds ratios (ORs) [34]. To demonstrate the effect of sector, we report results as predicted probabilities, which are often argued to be a more robust and intuitively interpretable effect size compared to ORs [35, 36]. Therefore, we will not report on the actual coefficients (although these can be found in appendix C2), but instead utilise the gologit2 output to compute non-linear probability models [31].

While marginal effects and predicted probabilities are not immune to unobserved heterogeneity [37], they are considered less sensitive to changes in the model specification than ORs [35]. All predicted probabilities are derived following marginal standardisation, i.e. as the average effect of sector on quality, as opposed to the effect of sector on quality on average (i.e. prediction at the means) [38]. We use the 'mtable' command to derive predicted probabilities, and the 'mchange' command to test if the changes in probability are statistically significant [39]. All differences in probabilities are reported in percentage points rather than as relative differences.

Explanatory variables

The key explanatory variables are sector, which will be coded as dummies: for-profit ("Private"), public ("Local Authorities"), and third sector ("Voluntary or Not for Profit"), with for-profit organisations as the benchmark category. We rely on the classification of service type provided by the CI and the analysis thus rests on the assumption that similarly classified service types do not differ systematically across sector. However, to account for the possibility that the effect of sector varies according to different services, all service types will be included as dummy variables. We exclude service types with very limited sector variation (i.e. adoption services, adult placement services, child minding, fostering services, nurse agencies, and offender accommodation services). We also ran the models on the largest sub-categories of services to explore generalisability across included service types (i.e. care home services, children day care services, housing support services, and support services). To ensure that observed effects were not driven by experience of service providers, we control for the number of years of registration with the CI. Further, we include dummy variables for whether an organisation has improved, declined or stagnated in its lowest ranked category since the last inspection. Finally, we control for the client group of care organisations (e.g. children, older people, dementia patients), whether an organisation is currently active, and the number of employed staff.

Dependent variables

The dependent variables of interest are the quality and risk domains described in Table 2. To test if sector predicts variation in regulation, we also investigate if organisations from different sectors vary in their probability to having been issued a requirement as result of an inspection. Further, to ensure that the effect of sector on quality aligns to a user-reported outcome, we explore the probability of organisations having been subject to at least one complaint within the last three years.

Due to a limited number of observations in each group, we merged category 1 and 2 ('poor quality'), category 3 and 4 ('adequate quality'), and category 5 and 6 ('good quality') across all quality domains. However, all reported analyses were also run on the original quality scores, which demonstrate similar results. Number of requirements and complaints were treated as binary variables (having received a requirement/complaint or not within last three years).

Patient and public involvement

Patients and public were not involved.

Results

Descriptive statistics

The dataset includes 8,111 (60.94%) for-profit organisations (FPOs), 2,602 (19,55%) organisations from the third sector (TSOs), and 2,597 (19.51%) public sector organisations (PSOs). Table 3 shows that FPOs have a higher number of employed staff on average, compared to PSOs and TSOs, but that the spread of observations for staff in FPOs is also the highest. TSOs have the highest proportion of organisations (4.66%) which improved their minimum quality rating within the last month, although the difference between sectors is modest. PSOs have, on average, been registered the longest, whereas FPOs tend to have been registered for a shorter period than PSOs and TSOs.

Table 3: Descriptive characteristics. Percentages in parenthesis for number of organisations, active organisations and changes in minimum quality ratings. Standard deviations in parenthesis for the average number of staff and years of registration.

	For-profit sector	**Public sector**	**Third sector**
Number of organisations	8,111 (60.94 %)	2,597 (19.51 %)	2,602 (19.55 %)
Active organisations			
Active	7,891 (97.29 %)	2,572 (99.04)	2,566 (98.62 %)
Inactive	220 (2.71 %)	25 (0.96 %)	36 (1.38 %)
Change in minimum quality rating in the last month			
Same	6,576 (93.44%)	2,363 (93.14%)	2,246 (91.08%)
Decline	215 (3.05%)	96 (3.78%)	105 (4.26%)
Improve	247 (3.51%)	78 (3.07%)	115 (4.66%)
Average number of staff	14.56 (31.87)	10.62 (27.27)	12.76 (24.68)
Average years of registration	8.58 (5.42)	13.92 (3.77)	11.53 (4.90)

Table 4 displays the cross tabulations of whether organisations have been subject to requirements after an inspection and if any complaints have been posed towards an organisation within the last 3 years. We see that PSOs received the fewest (proportionally) requirements (13.28%), and that FPOs and TSOs receive similar numbers of requirements (18.31% vs 19.43%). FPOs received a higher proportion of complaints compared to PSOs and TSOs, although the difference between the for-profit and third sector is relatively small (10.84% vs 8.65%).

Table 4 further shows that there is a roughly equal proportion of TSOs and FPOs with a 'poor' quality rating, and that PSOs tend to have the lowest proportion of organisations in this category. However, more FPOs score 'adequately' relative to PSOs and TSOs, and, conversely, a higher proportion of PSOs and TSOs fall within the high-quality category, compared to FPOs. One exception is the 'quality of environment' domain, in which the proportions across different levels of quality is roughly similar across sector. PSOs have the highest proportion of 'low risk'

organisations, whereas TSOs seem to be of proportionally higher risk compared to public and for-profit organisations.

Table 4: Cross tabulation of sector over quality domains, risk score, and requirements/complaints posed after inspection within last three years. Percentages in parentheses.

	For-profit sector	**Public sector**	**Third sector**
Quality of care (n=12,322)			
Poor quality	57	10	20
	(0.78)	(0.39)	(0.80)
Adequate quality	2,641	732	718
	(36.31)	(28.73)	(28.71)
Good quality	4,575	1,806	1,763
	(62.90)	(70.88)	(70.49)
Quality of staff (n=7,933)			
Poor quality	59	8	25
	(2.04)	(0.31)	(1.00)
Adequate quality	1,475	783	875
	(51.09)	(30.74)	(35.01)
Good quality	1,353	1,756	1,599
	(46.87)	(68.94)	(63.99)
Quality of environment (n=10,293)			
Poor quality	32	7	11
	(0.48)	(0.32)	(0.81)
Adequate quality	2,515	759	553
	(37.58)	(35.60)	(40.84)
Good quality	4,145	1,408	790
	(61.94)	(64.77)	(58.35)
Quality of management (n=12,386)			
Poor quality	107	15	38
	(1.47)	(0.59)	(1.52)
Adequate quality	3,642	1,041	1,033
	(50.12)	(40.87)	(41.83)
Good quality	3,518	1,491	1,441
	(48.41)	(58.54)	(56.65)
Risk of services (n=13,268)			
Low risk	7,263	2,420	2,259
	(89.96)	(93.22)	(86.95)
Medium risk	504	140	231
	(6.24)	(5.39)	(8.89)
High risk	307	36	108
	(3.8)	(1.39)	(4.16)
Complaints received within last 3 years			
No	7,233	2,453	2,377
	(89.16)	(94.46)	(91.35)
Yes	879	144	225
	(10.84)	(5.54)	(8.65)
Total	8,112	2,597	2,602
Requirements posed after inspection within last 3 years			
No	4,944	2,025	1,737
	(81.69)	(86.72)	(80.57)
Yes	1,108	310	419
	(18.31)	(13.28)	(19.43)
Total	6,052	2,335	2,156

Predicted probabilities

Figure 1 shows the predicted probabilities of all outcome variables (the 4 quality domains, risk, requirements, and complaints) across sector. The probabilities are derived from the generalised ordered logistic regression models, controlling for years of registration, type of sector, whether an organisation is active, target group of service providers (e.g. dementia), type of service provider, number of employed staff, and whether an organisation has improved, declined or stagnated in its lowest ranked quality category since the last inspection. The number of included observations in the generalised logistic models are lower than in Table 5 due to the exclusion of service types with limited sector variation (i.e. child-minding, adoption services, adult placement services, nurse agencies, and offender accommodation services). However, the observed effect of sector on quality remains similar when running more simple and inclusive models. For the binary outcome variables (complaints and requirements), the predicted probabilities are derived from multiple logistic regression models, controlling for the same variables as described above. To adjust for multiplicity, all displayed confidence intervals are Bonferroni corrected.

Figure 1: Predicted probabilities for quality (care, staff, management, and environment), risk, complaints, and requirements across sectors

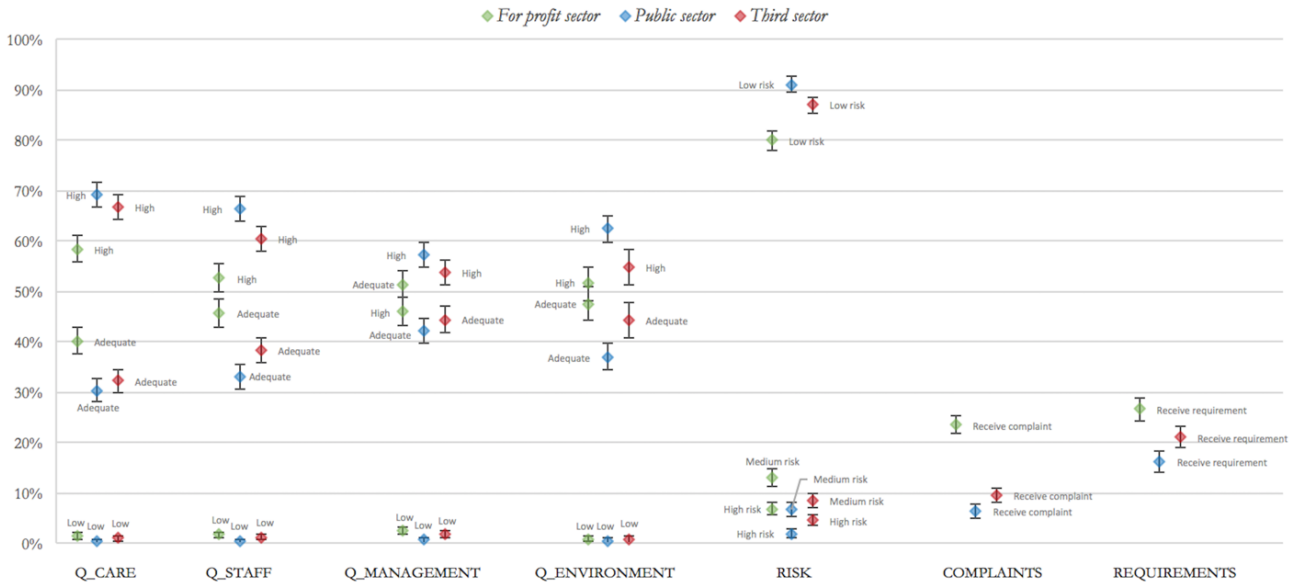


Figure 3: Predicted probabilities for the 4 quality domains, risk, complaints, and requirements across sector with 95% confidence intervals. All quality outcomes are based on a scale of low, adequate, and high quality. The risk outcome is based on a scale of low, medium, and high risk. The complaint and requirement outcomes are binary with the predicted probabilities referring to the likelihood of having been issued a complaint or requirement within the last three years. All predicted probabilities are displayed in percentages.

Difference between for-profit, public and third sector organisations

Figure 1 shows a similar pattern to the cross tabulations of quality across sector (see tables A3-A4 in appendix C1 for predicted probabilities). Again, FPOs are, in general, more likely to secure 'adequate' quality scores, but less likely to be in the 'high' quality category, relative to public and third sector organisations. For example, the predicted probability for FPOs to be in the 'high quality' category in the care and support domain is 58.4%, whereas it is 69.2% for PSOs, and 66.8% for TSOs. The predicted probability for FPOs to be of 'adequate quality' in the same domain is 40.1% vs 30.3% for PSOs and 32.2% for TSOs.

In table 5, we see that the differences in predicted probabilities for 'adequate' and 'high' quality are mostly statistically significant for public and third sector providers, compared to FPOs. The only quality domain which did not display this pattern was 'quality of environment', in which the differences in predicted probabilities between FPOs and TSOs were not statistically significant.

Further, FPOs were more likely to given a 'poor quality' rating compared to PSOs for 'quality of care', 'quality of staff', and 'quality of management', although the absolute differences are relatively small (below 2 percentage points). There were no statistically significant differences in the 'poor quality' rating between TSOs and FPOs.

Differences between public and third sector providers

We observe several significant differences in the probabilities between public and TSOs, although the differences in probability are smaller compared to the for-profit sector. For 'quality of staff', quality of environment, and 'quality of management, PSOs were more likely to be in the high-quality category compared to TSOs (66.4% vs 60.5%, 62.5% vs 54.8%, and 56.2% vs 53.7%, respectively). Similarly, TSOs were more likely to be rated of 'adequate quality' in 'quality of staff' (38.3%) and 'quality of environment' (44.2%) relative to PSOs (33.72 and 37%, respectively). For 'quality of care', 'quality of staff' and 'quality of management', TSOs are slightly, but statistically significantly, more likely to be in the low-quality group relative to PSOs (difference below 1.5 percentage points).

Risk of services

For-profit providers were the least likely to be classified as 'low risk' services (79.9%), and the most likely to be in 'medium risk' (13.2%) and 'high risk' (6.9%) groups. PSOs had the highest probability of being categorised as 'low risk' services (91.1%), and the lowest probability of having their services classified as of 'medium' (6.9%) and 'high risk' (2%), followed by TSOs (86%, 8.5%, and 4.5%, respectively). Most differences across the levels of risk between sectors were statistically significant, but the differences between PSO and TSOs were substantially smaller compared to TSOs vs FPOs, and PSOs vs FPOs.

Table 5: This table displays the difference in predicted probabilities across sector over the quality and risk domains with 95% confidence intervals.

	Low quality	Adequate quality	High quality
Quality of care and support (n=7,011)			
Public vs for-profit sector	-0.011***	-0.098***	0.108***
95% confidence interval	-0.018 to -0.004	-0.135 to -0.061	0.071 to 0.146
Third vs for-profit sector	-0.005	-0.079***	0.084***
95% confidence interval	-0.014 to 0.003	-0.116 to -0.042	0.047 to 0.121
Third vs public sector	0.006	0.019	-0.024
95% confidence interval	-0.001 to 0.012	-0.015 to 0.053	0.010 to 0.261
Quality of staff (n=7,007)			
Public vs for-profit sector	-0.013***	-0.125***	0.138***
95% confidence interval	-0.020 to -0.006	-0.164 to -0.087	0.100 to 0.177
Third vs for-profit sector	-0.005	-0.074***	0.079***
95% confidence interval	-0.014 to 0.004	-0.112 to -0.036	0.041 to 0.117
Third vs public sector	0.008*	0.051*	-0.059***
95% confidence interval	0.002 to 0.015	0.015 to 0.086	-0.095 to -0.023
Quality of environment (n=5,099)			
Public vs for-profit sector	-0.003	-0.106***	0.110***
95% confidence interval	-0.010 to 0.004	-0.151 to -0.061	0.065 to 0.154
Third vs for-profit sector	0.001	-0.034	0.033
95% confidence interval	0.004 to 0.010	-0.082 to 0.015	0.015 to 0.081
Third vs public sector	0.004	0.073***	-0.077***
95% confidence interval	-0.004 to 0.012	0.030 to 0.115	-0.119 to -0.034
Quality of management (n=7,004)			
Public vs for-profit sector	-0.019***	-0.093***	0.112***
95% confidence interval	-0.028 to -0.010	-0.133 to -0.133	0.073 to 0.152
Third vs for-profit sector	-0.007	-0.070***	0.077***
95% confidence interval	-0.018 to 0.004	-0.109 to -0.031	0.039 to 0.116
Third vs public sector	0.012***	0.023	-0.035
95% confidence interval	0.004 to 0.020	-0.014 to 0.060	-0.072 to 0.002
Risk of services (n=7,335)	**Low risk**	**Medium risk**	**High risk**
Public vs for-profit sector	0.112***	-0.063***	-0.049***
95% confidence interval	0.086 to 0.138	0.086 to -0.040	-0.064 to -0.034
Third vs for-profit sector	0.071***	-0.047***	-0.024*
95% confidence interval	0.045 to 0.098	-0.071 to -0.024	-0.041 to -0.007
Third vs public sector	-0.041***	0.016	0.025***
95% confidence interval	-0.063 to -0.018	-0.004 to 0.036	0.012 to 0.038

Note: * = p<0.05. *** = p<0.001. Also, higher values of risk indicate higher risk of services.

.

Complaints and requirements

Figure 1 shows the predicted probabilities for organisations to have been subject to a complaint or requirement in the last 3 years across sector. Again, we observe that FPOs are significantly

more likely to having been issued at least one complaint, compared to PSOs and TSOs, with a difference of 17.1 and 14 percentage points, respectively (see Table 6). Similarly, FPOs were more likely to having been subject to a requirement after inspection, relative to PSOs and TSOs with a difference of 10.5 percentage points for PSOs and FPOs, and 5.6 percentage points for TSOs and FPOs. PSOs were significantly more likely to having received a requirement or complaint relative to TSOs, but the differences were substantially smaller compared to FPOs.

Table 6: This table displays differences in predicted probabilities to have been issued a complaint or requirement in the last three years across sector.

	Difference in predicted probability across sector	95 % CIs
If a complaint has been issued within last 3 years (n=7,358)		
Public vs for-profit sector	0.171***	0.197 to 0.144
Third vs for-profit sector	0.140***	0.167 to 0.112
Third vs public sector	-0.031***	-0.012 to -0.050
If a requirement has been issued after inspection within last 3 years (n=6,320)		
Public vs for-profit sector	0.105***	0.138 to 0.072
Third vs for-profit sector	0.056***	0.138 to 0.072
Third vs public sector	-0.049***	-0.019 to -0.079

Note: * = p<0.05. *** = p<0.001

Other explanatory factors

The number of years an organisation has been registered with the CI and the number of employed staff were both significant predictors of quality. The association between age and quality showed that the older the organisation, the higher probability of being rated of high quality. Further, the more paid staff an organisation had, the more likely an organisation was to be in the 'adequate' category, whereas the probability of being in the 'high' quality category decreased with the number of employed staff. Neither the effects of staff and time in operation on quality were interacted by sector.

Sub-analyses

We reran the analyses on the biggest sub-groups of services (i.e. care homes, support services, child care services, and housing support services), which can be found in in table A5 in Appendix C1 and figures B1-B5 in Appendix C3. Most of these sub-analyses resulted in slightly

reduced effects of sector and with less clear differences between PSOs and TSOs. However, both the raw cross tabulations and the predicted probabilities generally display the same substantive pattern and direction of effects with FPOs being more likely to be classified as of adequate quality and of higher risk relative to TSOs and PSOs, which are more likely to be rated as high quality and of lower risk.

Discussion

The present study finds that public and third sector providers are more likely to be rated of high quality and of lower risk compared to FPOs, which are more likely to be rated of adequate quality and of higher risk. This pattern was similarly observed in FPOs being more likely to receive requirements and complaints compared to their public and third sector counterparts.

The results can be thought to correspond with the notion that TSOs and PSOs are primarily driven by pro-social values, and will thus be more motivated to maximise quality of services. Conversely, it may be contended that rational FPOs, being primarily motivated by profits, will coordinate their resources so that services fall within adequate standards. While PSOs were often less likely to be rated of 'low quality' relative to TSOs and FPOs, sector was generally not a strong predictor of the likelihood of organisations being given 'low quality' rating, which may be understood as all sectors being similarly incentivised not to be rated below acceptable threshold.

These findings echo a recent study on the quality of nursing homes, which utilised data from the Care Quality Commission (CQC) in England and Wales to test whether quality ratings varied according to sector finding that third and public providers were significantly more likely be rated of higher quality than FPOs, as measured by the quality index of the CQC [40]. However, contrary to our findings, the paper did not identify any statistically significant differences between third and public sector nursing homes.

Strengths and limitations

This study adds to the existing research in several ways. First, the present study replicated results of past research in a new context controlling for more factors (e.g. number of staff) and on a larger range of outcomes (e.g user-reported complaints) than in prior research. Second, the effect of sector on quality were robust to subsequent sub-analyses and may thus be thought to generalise across the included types of care services in the Scottish context. Further, contrary to existing research, the study identified several significant differences between PSOs and TSOs, which often favoured the performance of PSOs. Yet, these differences in quality were substantially smaller compared to for-profit vs public and third sector providers.

This study is subject to several limitations, some of which are due to the nature of the data. For example, the lack of effect of sector on quality for levels of low quality (1-2) may be driven by the low number of observations in these categories. However, that origin of sector had no effect on quality is consistent to the idea that organisations from different sectors will, theoretically, have a similar disincentive to be rated as inadequate.

A more substantial limitation of the analysis is that service types may differ systematically across sectors. For example, it may be that clients served by FPOs are inherently different from those served by third and public sector providers. To address this issue, we controlled for all organisational characteristics allowed for by the data (e.g. service types, years of registration, and number of employees), and ran the models separately on all major service types, showing the same substantive pattern as reported in the main results. Although all models adjusted for the types of clients served by the social care providers (e.g. children, older people, dementia patients), the data collected by the Care Inspectorate do not include individual level information on clients served, and we could thus not investigate client variation in terms of, for example,

socio-economic status across sectors. To the best of our knowledge, there is no existing research investigating client variation across sectors for social care services [3, 40, 41], which is important to explore in future research.

The analysis is built on the assumption that the quality ratings by the CI can be considered reliable indicators for the latent construct of 'quality', which may be a flawed assumption. Yet, the observed effects were consistent across different quality domains, which suggest that the effects of sector were - at least - robust to the construct of quality employed by the CI. Also, the pattern observed for the quality and risk domains was consistent with the inspection-requirements and complaints outcomes.

Last, the results should be interpreted in the light of the context of the data. The analysis was conducted on Scottish data on certain types of care services and the findings therefore primarily apply to that context. To test whether the observed effects do indeed transfer to other countries and service settings (e.g. healthcare), further research must replicate these analyses using comparable outcomes.

Implications for policy and practice

Research on the typology of sector tend to be based on theory with limited attention on what the features distinct to each sector mean in practice, and whether they influence aspects such as service delivery. Importantly, the findings of this paper do not indicate that services by FPOs are poor; in fact, the findings demonstrate that origin of sector was, for the most part, not a strong predictor of whether care organisations performed below acceptable standards. Also, similar to past research, the findings do not imply a causal effect of sector due to the observational nature of the data [42]. However, the findings can be interpreted to imply that care organisations driven by pro-social motivations tend to perform better than those that are not, although this remains a

theoretical interpretation, as the analysis does not include a concrete measure of motive. Going forward, it is essential to explore *mechanisms* to investigate whether the association found in this analysis is spurious or indeed the result of some sector-specific traits.

Conclusion

This paper investigated if origin of sector influence the quality of social care services regulated by the CI in Scotland to find that providers from the third and public sectors were rated of higher quality and lower risk than for-profit providers. Importantly, the analysis does not provide any empirical answer to *why* differences across sector are observed. Although the results are consistent with theory and past research, more research is needed to explore what exact characteristics account for the variation in quality observed across sectors. From the findings, it follows that regulators and policy-makers should incorporate more thought on how organisations across different sectors are incentivised to manage their resource, as this might influence the subsequent performance of commissioned services. However, there is little research addressing how this is best realised in practice, which warrant further analysis.

Availability of data and material: All data generated or analysed during this study is publically available or attached in the supplementary information files.

Author contributions: The research design and objectives were developed by ABM and PM. ABM designed and conducted the statistical analysis and drafted the manuscript with input from PM. All authors read and approved the final manuscript.

Acknowledgements: The authors would like to express their gratitude to Caspar Kaiser, Bastian Betthäuser, Jonathan Klüser, Ani Movsisyan, Thees Spreckelsen, Peter Bach-Mortensen, and Professor Kimberly Scharf for their invaluable feedback on earlier drafts of the manuscript. We

would also like to thank the reviewers for their helpful comments, which improved the manuscript substantially.

Bibliography

1. Dickinson H, Allen K, Alcock P, Macmillan R, Glasby J. The Role of the Third Sector in Delivering Social Care. NIHR School for Social Care Research. 2012. doi:978-0-85328-443-7.
2. Macmillan R. The third sector delivering public services: an evidence review. TSRC Working Paper 20 (Birmingham, University of Birmingham). 2010. https://www.birmingham.ac.uk/generic/tsrc/research/service-delivery/wp-20-public-services-evidence-review.aspx. Accessed 7 Apr 2016.
3. Comondore VR, Devereaux PJ, Zhou Q, Stone SB, Busse JW, Ravindran NC, et al. Quality of care in for-profit and not-for-profit nursing homes: systematic review and meta-analysis. Bmj. 2009;339 aug04 2:27–32. doi:10.1136/bmj.b2732.
4. Tynkkynen L-K, Vrangbæk K. Comparing public and private providers: a scoping review of hospital services in Europe. BMC Health Serv Res. 2018;18:141. doi:10.1186/s12913-018-2953-9.
5. Grand J Le, Le Grand J. Quasi-Markets and Social Policy. Econ J. 1991;101 September:1256–67. doi:10.2307/2234441.
6. Le Grand J. Quasi-Market versus State Provision of Public Services : Some Ethical Considerations. Public Reason. 2011;3:80–9.
7. Almeida Á. The Role of Private Non-For Profit Healthcare Organisations in the NHS system: Implications for the Portuguese Hospital Devolution Program. 2016;577. http://wps.fep.up.pt/wps/wp577.pdf. Accessed 11 Jan 2018.
8. Pritchett L. It pays to be ignorant: A simple political economy of rigorous program evaluation. J Policy Reform. 2002;5:251–69. doi:10.1080/1384128032000096832.
9. Chapman T, Bell V, Robinson F. Measuring impact: easy to say, hard to do: a think-piece to stimulate sector debate from the Third Sector Trends Study. 2011;:17.
10. Harlock J. From outcomes-based commissioning to social value? Implications for performance managing the third sector. TSRC Working Paper 123 (Birmingham, University of Birmingham). 2014. https://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-123.pdf. Accessed 11 Dec 2017.
11. Titmuss RM. The Gift Relationship: From Human Blood to Social. London: George Alien & Unwin Ltd.; 1970. https://www.cabdirect.org/cabdirect/abstract/19712702984. Accessed 25 Jan 2018.
12. Hardwick R, Anderson R, Cooper C. How do third sector organisations use research and other knowledge? A systematic scoping review. Implement Sci. 2015;10:84. doi:10.1186/s13012-015-0265-6.
13. Moore MH. Managing for Value: Organizational Strategy in for-Profit, Nonprofit, and Governmental Organizations. Nonprofit Volunt Sect Q. 2000;29:183–204. doi:10.1177/089976400773746391.
14. Anheier HK. Nonprofit Organizations: Theory, management, policy. Routledge; 2005. http://books.google.com/books?id=s_uzg-KoVyIC&pgis=1.
15. Alcock P. A strategic unity: defining the third sector in the UK. Volunt Sect Rev. 2010;1:5–

24.
16. Macmillan R. The third sector delivering public services: an evidence review. Policy Press; 2010.
17. Witesman EM, Fernandez S. Government Contracts With Private Organizations: Are There Differences Between Nonprofits and For-profits? Nonprofit Volunt Sect Q. 2013;42:689–715. doi:10.1177/0899764012442592.
18. Salamon LM, Anheier HK. In search of the non-profit sector. I: The question of definitions. Voluntas. 1992;3:125–51. doi:10.1007/BF01397770.
19. Herrera CA, Rada G, Kuhn-Barrientos L, Barrios X. Does Ownership Matter? An Overview of Systematic Reviews of the Performance of Private For-Profit, Private Not-For-Profit and Public Healthcare Providers. PLoS One. 2014. doi:10.1371/journal.pone.0093456.
20. Kelly J. Reforming public services in the UK: Bringing in the third sector. Public Adm. 2007;85:1003–22. doi:10.1111/j.1467-9299.2007.00680.x.
21. Cabinet Office. Open Public Services: White Paper. 2011. http://www.cabinetoffice.gov.uk/sites/default/files/resources/open-public-services-white-paper.pdf%5Cnhttp://www.openpublicservices.cabinetoffice.gov.uk/.
22. Kramer RM. A Third Sector in the Third Millennium? Int J Volunt Nonprofit Organ. 2000;11.
23. Datastore. http://www.careinspectorate.com/index.php/statistics-and-analysis/data-and-analysis. Accessed 26 Jan 2018.
24. New Care Standards. New Care Standards | Review of Scotland's National Care Standards. 2017. http://www.newcarestandards.scot/. Accessed 30 Aug 2017.
25. Care Inspectorate. Risk Assessment – Information for Care Service Providers. 2017. http://www.careinspectorate.com/images/documents/148/Risk Assessment – Information for Care Service Providers.pdf. Accessed 28 Aug 2017.
26. Care Inspectorate. How we inspect: What goes into our inspection reports for care services. 2013.
27. Inspectorate C. Inspections. http://www.careinspectorate.com/index.php/inspections. Accessed 9 Feb 2018.
28. Care Inspectorate. Inspection Handbook. 2016;109 August.
29. Vandenbroucke JP, Von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration , for the STROBE Initiative. 2014.
30. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for reporting observational studies. Prev Med (Baltim). 2007;45:247–51.
31. Williams R. The Journal of Mathematical Sociology Understanding and interpreting generalized ordered logit models Understanding and interpreting generalized ordered logit models. J Math Sociol. 2016;40:7–20. doi:10.1080/0022250X.2015.1112384.
32. Brant R. Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. Biometrics. 1990;46:1171. doi:10.2307/2532457.
33. Williams R. gologit2 documentation. Stata J. 2006;6. http://www.nd.edu/~rwilliam/gologit2/gologit2.pdf. Accessed 29 Jan 2018.
34. Mood C. Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It. doi:10.1093/esr/jcp006.

35. Norton EC, Dowd BE. Log Odds and the Interpretation of Logit Models. Health Serv Res. 2017;:1–20.
36. Norton EC. Log Odds and Ends. 2012. doi:10.3386/w18252.
37. Skopek J. One Hammer, Different Nails – A Note on the Confusing Sociologists ' Debate on Comparing Coefficients in Logistic Regression. 2015. doi:10.13140/RG.2.1.2181.7846.
38. Muller CJ, Maclehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. Int J Epidemiol. 2014;:962–70. doi:10.1093/ije/dyu029.
39. Long JS, Freese J. Regression models for categorical dependent variables using Stata. 2014. http://ideas.repec.org/b/tsj/spbook/long2.html.
40. Barron DN, West E. The quasi-market for adult residential care in the UK: Do for-profit, not-for-profit or public sector residential care and nursing homes provide better quality care? Soc Sci Med. 2017;179:137–46. doi:10.1016/j.socscimed.2017.02.037.
41. Hjelmar U, Bhatti Y, Petersen OH, Rostgaard T, Vrangbaek K. Public/private ownership and quality of care: Evidence from Danish nursing homes The involvement of private for-profit (FP) and not-for-profit (NFP) providers in the. 2018. doi:10.1016/j.socscimed.2018.09.029.
42. Konetzka RT. Do not-for-profit nursing homes provide better quality? BMJ. 2009;339 aug04 2:b2683–b2683. doi:10.1136/bmj.b2683.

Chapter 6: Discussion and recommendations

Summary of findings

This thesis has explored factors influencing the ability of TSOs to evaluate and implement their services, and examined quality variation among Scottish social care providers. In so doing, this thesis has systematically, identified, appraised, and synthesised a broad range of research, as well as analysed administrative data to meet its research objectives. The findings can be summarised as follows:

Paper 1: Barriers and facilitators to evaluation among third sector organisations

This study, presented in Chapter 3, set out to investigate the factors reported to influence the ability of third sector practitioners to evaluate their work. The systematic review identified, appraised, and synthesised a range a primary research on the factors reported to influence the ability of TSOs to evaluate their work. It found that lack of resources and evaluation skills were the most commonly reported barriers to evaluation. Further, practitioners often reported struggling with identifying relevant evaluation metrics that could be meaningfully applied to their work. The main facilitating factors included evaluation support, an organisational culture embracing evaluation, and having accountability structures in place. These findings were consistent across study quality and study type.

Importantly, the findings indicate that TSOs often report being dissatisfied with the appropriateness and usefulness of the reporting requirements imposed on them by funders and commissioners. Specifically, TSOs often struggled with the funding requirements and also reported to be forced, as a requirement of their commissioning, to evaluate their services without getting the necessary support and training to do so. These findings may be utilised to improve the commissioning process going forward. For example, funders may use these results to

consider how they can incorporate evaluation as part of the funding to reduce the risk of it being a bureaucratic burden to TSOs. Asking TSOs to evaluate their work as part of funding is important for the accountability of the work, and it is therefore important that this evaluation is undertaken reliably. Further, a main theme was the challenge to identify relevant outcome frameworks to which TSOs could adhere. In ensuring that the evaluation criteria imposed on TSOs are relevant to practice, funders ought to include practitioners when deciding what outputs and outcomes to focus on.

Paper 2: Barriers and facilitators to implementing evidence based interventions among third sector organisations

This study which was presented in Chapter 4 focused on identifying the factors affecting the ability of TSOs to implement evidence-based interventions (EBIs). Using rigorous best-practice systematic review methodology, the review synthesised 31 primary studies reporting on the implementation process of TSOs. The review found that the main reported themes regarding barriers were: (1) lack of resources, (2) recruitment and client-resistance issues, (3) lack of implementation expertise, and (4) divergence between the EBI and organisational values. The most common themes of reported facilitators were (1) congruence between organisational culture and the EBI, (2) implementation expertise and support, and (3) motivation to deliver interventions that 'work'. The synthesis identified that poor guidance and/or flexibility around *adaptation* of the EBIs (e.g. to suit the needs of a population different than the one the EBI was designed for) was a central problem to many practitioners.

These findings have important implications for existing practices. Through a comprehensive and rigorous appraisal of existing research, the review identified that practitioners implementing EBIs were often uncertain about what features of it could be customised in the implementation process – and which could not. This issue was driven by a lack of clarity on the 'core contents'

or 'active ingredients', i.e. the intervention components thought central to drive the intervention effect. In fact, EBIs are often offered in 'black-box' one-size-fits-all packages, which do not clearly specify the intervention logic [1], thus making it unclear how to adapt the intervention without compromising important aspects. Conversely, the findings can also be interpreted to suggest that funders ought to evaluate the fit of an EBI and the TSO supposed to implement the intervention more carefully before commissioning the intervention. Thus, commissioners and EBI developers should make efforts and provide guidance on how to optimise the match of the implementing organisation and the EBI.

Paper 3: Quality variation among for-profit, public, and third sector social care providers

This final study which was presented in Chapter 5 investigated variation in quality related outcomes among for-profit, public, and third sector social service providers in Scotland. Including a number of social provider types, such as but not limited to, nursing homes, children homes, and housing support services, the paper identified statistically significant variation across different types of ownership. All included service types were subject to the same regulatory framework (i.e. *the social care and improvement regulations Scotland* [2]) and regulated by the Care Inspectorate in Scotland. The analysis demonstrated that third and public sector providers were generally more likely to be rated of very high quality, whereas for-profit were more likely to be rated of 'adequate' quality. This pattern was consistent across most quality indicators and also illustrated by PSOs and TSOs being rated, on average, to be providing lower risk services than their for-profit counterparts. The findings were robust to controlling for a number of organisational characteristics and also proved consistent across all included service types.

It is important to stress that this analysis does not suggest that for-profit providers are performing poorly; in fact, for levels of 'low' quality and high risk, there were little to no variation among for-profit, public, and third sector providers. However, the analysis does suggest

that public and third sector social care providers are, on average, more likely be rated of high quality and of low risk – as assessed by the Care Inspectorate – compared to for-profit providers. These findings constitute an important resource for regulators, in that the paper offers a rigorous analysis mapping systematic variation in social service providers across ownership. This finding is of particular significance considering the swiftly changing social care landscape, which is increasingly dominated by for-profit providers as a result of the quasi-market commissioning structure. However, in discussing whether and how this variation can, or should, be met through regulation, more research elaborating on the specific issues across ownership is required.

Limitations

This thesis is subject to a number of limitations, some related to the methodology of the individual chapters and others due to the nature of the questions posed by this thesis. These are addressed below.

Paper specific limitations

Searching for non-profit literature

A fundamental precondition of a successful systematic reviews is the ability to identify and retrieve the studies relevant to the posed questions. A systematic review is, by definition, constrained to the research it includes and flaws in the search strategy are thus critical. However, no gold-standard robustness measure exists to ensure that a search is relevant to the question it attempts to answer. Therefore, it is critical that a search is transparent and reproducible. One of the main aims of papers 1 & 2 [3, 4] was to connect the non-profit, public administration, and organisational literature to health services research. The focus of the search strategy of these reviews were therefore on *sensitivity*, and the search strings were kept broad and with consideration to the variable use of terms like 'non-profits'. The University Social Science

librarian was consulted to avoid any obvious flaws in the search logic. Both systematic reviews were successful in identifying a range of multidisciplinary research, using very different terms for 'non-profit organisations', thus suggesting the search was successful in retrieving studies using variable terms for the same concept. Further, the reference lists of all included studies were manually examined to ensure no important citations were missed.

Synthesising barriers and facilitators

Papers 1 & 2 relied primarily on aggregative and descriptive synthesis techniques. Reviews of barriers and facilitators are often criticised for relying on descriptive and aggregative synthesis approaches (such as 'vote counting'), which are generally considered the least rigorous and often most biased forms of synthesis [5, 6]. Aggregative and descriptive "syntheses" are accused, among other things, of failing to account for the methodological quality of included studies in forming their interpretations, of ignoring the role of context and contingency in constructing synthetic findings, and of incorrectly conflating the frequency with which findings (e.g. barriers) appear in primary studies with the importance of those findings. Importantly, no guidelines exist on the appropriate conduct of these reviews, although this type of review is frequently published in high impact journals. However, to ensure that the reviews used rigorous practices that allowed robust representation of the included research, a number of measures were taken to avoid the common pitfalls of this type of review.

First, all studies were subject to rigorous quality appraisal to ensure that the findings were not driven by poor quality research. Specifically, the quality appraisal was used to inform the robustness and reliability of the findings. Second, all identified factors were thematically categorised rather than descriptively counted by frequency, thus accounting for factors circulating around similar underlying issues. Last, both reviews analysed the consistency with which each of the identified themes were reported on by the individual studies, and if study

quality and type (e.g. qualitative and interview studies) influenced the themes reported. This was done to ensure that findings were not driven by single studies or specific study characteristics. These efforts to improve existing reporting and synthesis practices in this type of review has informed a subsequent methodological paper on the conduct of barrier and facilitator reviews, which is currently in peer-review.

Limitations on defining 'quality'

Paper 3 [7] utilises the quality outcomes collected and operationalised by the Care Inspectorate in Scotland. Although these outcomes are constituted by a comprehensive and transparent checklist of items, these constructs may be flawed, in that 'quality' is a complex construct, which is not easily reduced to scores of 'low', 'medium', and 'high'. However, given that this scale is used for regulation purposes, adapting it for analysis entails some advantages in terms of how the results are understood by relevant stakeholders. By using the same measurements and quality constructs as the Care Inspectorate, the results can be easily interpreted by the practitioners and regulators to which they apply. Moreover, an ongoing project using the same dataset has manually coded the inspection reports from a full inspection year to find that the results reported in paper 3 similarly translate into likelihood to violate certain requirements and recommendations.

General limitations of thesis

Geographical relevance of findings

An overarching limitation to this research is the geographic relevance of the findings. While both reviews were not restricted to a geographical area, they include studies conducted in a variety of countries, although most were based in North America and Europe. On the one hand, this can be seen as an advantage, in that the reviews present a (relatively) global and generalisable picture of third sector practice, this inclusiveness could, on the other, be accused of overlooking important regionally embedded factors. Yet, given that little systematic research has been

conducted on these topics, it was important for this thesis to be inclusive, not just in terms of academic disciplines, but also in terms of the national context(s) of study. The current lack of definitional clarity on third sector research is arguably enhanced by restricting research on the third sector to single national contexts. While this approach may overlook important locally embedded nuances on TSOs, the only way to identify those is by doing inclusive and systematic research allowing for such comparisons.

Additionally, Chapter 5 used Scottish data and the findings therefore primarily apply to that context. However, it is important to note that the empirical analysis on which the paper is based was developed from a broad reading of the theoretical literature. All empirical work is, by definition, restricted to the empirical context of the investigation and the broader relevance thus depends on the plausible comparability of the setting investigated with that in question. Further, the Scottish context is arguably very similar to rest of the UK, particularly in terms of how social care services are regulated, which is, for example, demonstrated by the comparable direction in results identified in the English and Welsh context [8]. Thus, the findings may be thought to generalise to countries with similar regulation and policy structures.

Recommendations

The analyses conducted in this thesis have a number of implications that are highly relevant to existing practice and policy. These can be summarised as three specific recommendations, which are elaborated below.

Recommendation 1: Incentivise an evaluation culture of learning

To identify the main evaluation issues in the third sector, the first systematic review investigated the barriers and facilitators to evaluation among TSOs [3]. Overall, the review found that the

main factors preventing TSOs from evaluating their services was related to lack of finances, evaluation expertise and skills, and challenges regarding identifying and utilising existing evaluation and measurement frameworks. These challenges were often grounded in the increasing pressure to evaluate combined with the failure of funders and regulators to support practitioners in the evaluation process. Practitioners often reported that this mismatch between expectations and evaluation led to poor evaluation practice, and resulted in evaluation becoming a bureaucratic tick-box exercise with the goal of satisfying donors, as opposed to ensuring the safety and effectiveness of services.

To address these evaluation issues, it seems imperative to develop a common way to think about evaluation among practitioners, donors, service-users, and policy-makers, and centrally, to consider whether, and if so how, different types of services should be evaluated. Currently, there appears to be a divide between what TSOs are expected to deliver in terms of evaluation, and their ability to deliver it rigorously. Initiatives such as the Social Value Act [9] demonstrate that there exists a political will to enforce evaluation, which is obstructed by an absence of common language on 'impact' [10]. This challenge may be increased by the fact that many evaluation frameworks imposed on TSOs are arguably based on a 'New Public Management' line of thinking, in effect favouring practices from the private for-profit sector [11]. For example, commonly applied quality frameworks, such as ISO 9001 were originally developed in the for-profit sector for industrial organisations and emphasise aspects related to management and product maximisation [12, 13]. It has been argued that enforcing impact metrics originating from industrial practices influences the key characteristics of TSOs, forcing them to become more 'business-like' in order to compete for service contracts [14, 15].

While it may be considered positive that commissioners impose similar requirements for organisations across different sectors, it is important that funders and regulators consider what

type of *evaluation culture* they incentivise. For example, the pressure for TSOs to adhere to a new public management style of organisation is often criticized by third sector practitioners as being needlessly bureaucratic, and for enforcing criteria that are not relevant to practice [16, 17]. Importantly, many services delivered by TSOs, particularly those involving social intervention, are not appropriate for management quality systems designed for product maximisation. Adhering to such frameworks essentially becomes an exercise in satisfying funders/commissioners. For evaluation to be meaningful to practitioners, it is central that third sector practitioners acknowledge the appropriateness of the evaluation criteria they are subjected to, which necessitates stakeholder inclusive discussions on how criteria can best be customised to the nature of the services to which they apply.

Importantly, there are ethical and moral reasons to incentivise an appropriate evaluation culture, in addition to the potential increased impact and efficiency. Similar to the imperative *first do no harm* in medical practice, service providers – regardless of sector – delivering social interventions are ethically obliged to minimise the risk of harm, which require rigorous and continuous evaluation. This is partly the responsibility of practitioners, but given the tax-exempt status of TSOs and their dependence on funders, government or commissioners also hold a central responsibility in ensuring that TSOs are benefitting society and not causing harm. That is not to say that it should not be possible for local sports and community clubs to achieve 'charity' status and that all TSOs must report critically and rigorously on their activities. However, for organisations that aim to intervene (interfere) in the lives of vulnerable individuals, it befalls both practitioners, government, and commissioners to ensure that such services are safe.

Evaluation ought to be considered a management and safe-guarding tool that practitioners can utilise to monitor and improve existing practices. Yet, for evaluation to improve practice, it must be conducted following sound and rigorous methods using outcomes accepted by key

stakeholders. Thus, practitioners need to incorporate evaluation as an integral part of the service-delivery and abandon the practice of treating evaluation as a tick-box exercise and marketing/commissioning tool. To facilitate this development, it is important to have realistic and improvement-oriented expectations to service evaluation and acknowledge that evaluations might illuminate the fact that a service is, in fact, not effective.

Yet, given that positive results may directly influence the ability for TSOs to secure funding, it can be directly against management interests to conduct rigorous evaluations, and, in case of negative results, to publish the evaluation [18]. In the UK at least, TSOs, even those working with vulnerable population groups, are not *required* to publish their results. This problem contains parallels to *publication bias* in research, i.e. the issue of significant positive results being more likely to be published in high impact journals. There has been a range of initiatives to combat this issue in academia, including journals starting to accept null-findings and the development of 'best practice' reporting guidelines [19]. These initiatives constitute a step towards a research culture which acknowledges that 'null-findings' are important for practice and that acknowledges that rigorous methods do not imply significant results (and vice-versa) [20].

To reach consensus on what constitutes credible evaluation, more emphasis is warranted on the *evaluation process* as opposed to immediate outputs. For example, in intervention research there is a call for a *complexity* or *realist* perspective when evaluating social interventions with multiple interacting components [21]. This includes conducting process evaluations, which involve investigating the *implementation, context* and *mechanisms* of interventions, thus allowing for a better understanding of *how* and *under what conditions* a given intervention operates [22]. That is not to say that results are not central to decision-making, but to understand why an intervention works (or does not work), which is critical in assessing the potential for scale-up or dissemination, it is essential to consider aspects beyond intervention effects. Such a shift is no easy task and would

be disruptive to many existing practices in the third sector. Importantly, it necessitates a shift in *evaluation culture*, which places *learning* and *safe-guarding* as the main objectives of this work. Recommendations on how to improve current practices are summarised in Table 1.

Table 1: Recommendations on evaluation for funders and practitioners

Recommendations for funders	Recommendations for practitioners
– Funders and commissioners should include clear and stakeholder informed evaluation guidelines. – Funders should ear-mark resources for evaluation, including appropriate training, as part of larger grants. – The pursuit of magic bullets in alleviating social issues is appealing, yet risky for in-depth understanding of social interventions. Rigorous evaluation requires time, commitment, and might be inversely associated with effect size. Thus, evaluation efforts should be less concerned with identifying miracle interventions, and more concerned with understanding mechanisms and improving future practices.	– Clearly define the outputs and outcomes of a programme or intervention ex-ante. For complex social interventions without existing or well-established outcomes, this is commonly done in dialogue with patients, experts and other key stakeholders. – Define the evaluation method and time-line prior to implementation. This is important, as there is a range of well-defined biases involved with ex-post evaluation designs, which ultimately jeopardise the validity of the evaluation results. – Consider and continuously evaluate the risk of harm of interventions and programmes. There are multiple examples of well-meaning social interventions that ultimately harmed their beneficiaries, including Scared-Straight and Cambridge-Summerville [23]. To address this, it has been suggested that academics and practitioners develop a 'dark logic model' to identify potential mechanisms through which harm may occur prior to implementing interventions [24].

To enable a practice- and learning-oriented evaluation culture, it is essential to be transparent around evaluation results, which involves admitting when a programme is *not* working as intended, which should subsequently be used to revise existing practices. Also, to facilitate an evaluation culture revolving around learning, commissioners could allocate more resources in the selection process of what organisations to fund and include capacity and evaluation training as a required part of the funding. However, budgeting for evaluation is not enough in isolation; practitioners also need to demonstrate *motivation* to evaluate. There are examples of this kind motivation; for example, the non-profit organisation, Evidence Aid, recently published the RCT results of scaling up their 'No Lean Season' programme which essentially proved non-effective (https://www.evidenceaction.org/blog-full/why-test-at-scale-no-lean-season/). Notably, this publication received positive attention in media outlets (e.g. https://www.vox.com/2018/11/29/18114585/poverty-charity-randomized-controlled-trial-

evidence-action), in which they were praised for their openness on the discouraging programme results.

Stakeholders must abandon practices which consider service-delivery and evaluation separate activities. Rigorous evaluations take time and require planning, resources, and commitment. However, prominent third sector evaluators, such as the Charity Commission and Charity Navigator, seem to neglect the value of understanding intervention effects in favour of focusing on superficial outputs and financial reporting. Going forward, it is essential that efforts are made to ensure that regulations and commissioning schemes are facilitative of improving *outcomes*, as opposed to be incentivising superficial and positive evaluations. This requires a much clearer language and terminology on evaluation and impact, which should be developed in dialogue with practitioners and clients with the overall goal of enabling evaluation to improve services and management.

Recommendation 2: Design clearly defined interventions that can be adapted to local contexts and populations

One of the main limitations of the evidence-based practice (EBP) movement is that it is unclear to what extent the principles underlying the movement have successfully improved existing practices. Indeed, there is little systematic research on how evidence-based practice mechanisms are best incorporated into different settings, which is a significant blind spot. Paper 2 addressed this issue by investigating the factors influencing the ability of TSOs to implement evidence-based interventions (EBIs) [4]. The paper found that many third sector practitioners reported to struggle with adapting the commissioned EBIs to their target population, due to multiple reasons, including but not limited to, poor adaptation guidance and lack of expertise/funding.

Importantly, repositories of EBIs (e.g. the 'What Works Clearinghouses' and Diffusion of Evidence-based Interventions (DEBIs)) are becoming an increasingly popular way of disseminating commissionable interventions. Although these efforts are built on the intuitive rationale that interventions associated with positive effects should be widely disseminated, it is often not understood why an EBI works. In fact, the term 'EBI' is often poorly defined and often relates to an intervention having shown statistically significant effect in an efficacy trial. Although there is no universally accepted definition of an EBI, a recent article in the Annual Review of Public Health defined EBIs as "programs, practices, policies, and guidelines with proven efficacy" [25]. Importantly, such a definition overlooks the role of 'mechanisms' and a conceptualisation of how, and under what conditions, a given intervention is supposed to work [26]. In turn, although EBIs are associated with positive effects, they do not, according to the above definition, include a causal model defining how the intervention is supposed to operate.

In light of the rather vague definition underlying EBIs, it is hardly surprising that TSOs struggle with implementing these in practice. In fact, although paper 2 accepted all studies that reported to have implemented an EBI, these studies rarely defined nor elaborated on the actual intervention being implemented. A way to address this issue may be to clarify the language surrounding EBIs, particularly in terms of how they are supposed to work. Complex social interventions, such as parenting programmes, often entail multiple components that can plausibly be thought to influence one another [27]. For this kind of intervention it is therefore not only possible, but likely, that the effect will be influenced by the individual context and the implementation process [28].

Intervention adaptation is already an important aspect of implementation research. However, although it is commonly acknowledged that adaptation matters, it is often unclear how to best go about it in practice. A recent systematic review examining the adaptation of public health

interventions found that only a third of the adaptation studies reviewed (n=42) were done following a guidance framework [29]. This suggests that the demand for adaptation is not met with the required support to undertake it.

One way to address this is to draw on the principles of *realist evaluation* [30]. These stipulate that to understand social programmes or interventions, reducing intervention outcomes to average effect sizes do not suffice. Rather, realist philosophy suggests that outcomes should be understood as the interaction of mechanisms and context. A realist reviewer thus strives to conceptualise the configuration of context, mechanism, and outcome (CMO) with the aim of understanding 'what works, 'for whom', and 'under what conditions' [31].

By incorporating realist principles into the design and development of EBIs, the mechanisms underlying interventions may be better understood, which could facilitate more focused implementations. Yet, it is important to stress that outcomes, particularly relating to harms or other iatrogenic effects, remain a central concern when evaluating interventions. Adaptation and implementation fidelity needs, however, to be a prespecified part of developing interventions, rather than an afterthought. Putting a clearer emphasis on mechanisms, rather than assuming interventions that have proven efficacious in one setting will naturally operate similarly in another, might resolve the common criticisms that EBIs are 'black boxes'. Doing so will be important for the following reasons:

First, by conceptualising mechanisms driving the effect of social interventions, practitioners are better enabled to adapt interventions without undermining 'core components' or 'active ingredients' [32]. Second, a quintessential aspect to intervention research is considering the risk of harm [23]. However, it is not always understood *why* observed harms occur and it has thus been suggested that developers ought to develop a 'dark logic model' as part of the intervention

development to conceptualise why an intervention can be thought to negatively influence its participants [24]. In turn, putting a stronger emphasis on mechanisms may help safeguard intervention beneficiaries. Last, a better understanding of mechanisms will help optimise interventions. Rather than evaluating the utility of interventions entirely according to the average effects, an analysis of the potential interaction between the implementation context and intervention mechanisms will provide much richer data regarding the ways in which the intervention should be disseminated going forward.

Recommendation 3: Policy and regulation should facilitate incentives that focus on quality rather than costs

With the growing marketisation of the provision of social services [33–35], it is central to understand if and how the profit status of providers can be thought to influence the operation of services, so that practitioners and policy-makers can address problematic sector-specific differences through regulation [36, 37]. However, with the increased blurring of sectors due to open bidding in social and health care provisions, as well as advances in pay-for-performance schemes and private-public partnerships, it is unclear to what extent the demarcation of ownership is a helpful lens through which to analyse providers [15, 38–41]. Although a range of existing research suggest that for-profit, public, and third sector providers entail fundamental differences that materialise in their operation [42], it is not clear (empirically) *why* the profit status of organisations should matter to their organisation.

There are several reasons to think *why* ownership should matter for the operation of social care services. For example, it is conventionally argued that the behaviour of for-profit organisations is driven by the pursuit of profit maximisation, which in case of flawed commissioning or market incentives may compel such organisations to favour profits above clients' needs and service quality. Conversely, public and third sector providers are often thought to be driven by pro-

social incentives, suggesting that these types of organisations are more responsive to client-needs and more motivated to 'do good' [43]. However, this is not to say that TSOs are immune to inverse commissioning incentives [44], and Chapter 3 identified several issues in third sector practices, particularly related to the (dis)incentives to conduct rigorous evaluation [3, 18].

One of the main voiced concerns regarding the privatisation of social care services is that open competition might result in reduced service quality for those who cannot pay for private services [45]. Contrary to this possibility, the results of Chapter 5 demonstrate that public and third sector providers are more likely to be rated as being of very high quality (as rated by the Care Inspectorate), although ownership did not explain differences in very low quality ratings. The findings demonstrated in Chapter 5 [7] and those echoed in similar work [8, 46] suggest a general trend showing that for-profit providers perform below, as measured by independent inspecting agencies, their public and third sector counterparts. However, despite these associations between 'quality' and ownership, there is still a poor understanding of the causal mechanisms driving these differences, and it is therefore unclear how regulators should act on these findings.

The theoretical framework outlined in Chapter 2 suggests a few ownership specific principles that policy-makers can consider when designing their commissioning structure. These can be summarised by the following points:

(1) The ability and quality of social services implemented by FPOs is largely dictated by the market and contracting requirements/regulations. It can thus be expected that the quality of for-profit providers of social services will be compromised if the regulation/contracting framework allows for opportunistic behaviour.

(2) PSOs deliver services that are required by law and the quality that is dictated by regulation. One can expect this type of provider to deliver services according to the political will to budget and prioritise these services.

(3) TSOs respond to market and state failures and deliver services in accordance to community and societal need. Third sector service provisions can thus be hypothesised to reflect the quality and availability (or lack thereof) from other providers.

The market for care is complex, partly due to the fact that 'quality of care' is a latent construct that is challenging to capture and, in effect, to optimise. This is further complicated by the predominant influence of new public management commissioning structures, and with the growing emphasis on costs, partly as a result of the ageing population [47]. However, by failing to make meaningful efforts to define and operationalise quality as part of the commissioning process, the risk of focusing on easily measurable outcomes such as costing is introduced. Importantly, a superficial incentive framework is risky in an open market setting, as it may facilitate a 'race to the bottom', in which providers are forced to design their services around easy-to-serve clients to win contracts. Although costing is an important component when making commissioning decisions, it is essential that quality of care is also considered.

Future research directions

There are still several outstanding issues in need of attention going forward. These are elaborated on below.

Further research on third sector evaluation

First, a central outstanding issue involves whether and how evaluation data from TSOs should be made available, and, if so, to whom? While transparency is central to good practice,

confidentiality must not be compromised, especially for organisations working with information on vulnerable populations. Second, it remains unclear whether TSOs can expected to evaluate their activities following rigorous and transparent practices. Previously it has been argued that one of the main flaws of current practices is that TSOs utilise evaluation as a marketing and commissioning tool, which only leaves room for positive results [18]. We contend that commissioners and regulators hold a central role in shifting the current incentive framework to make it necessary and preferably attractive for third sector practitioners to conduct rigorous evaluations to promote more rigorous management towards the greatest effects for the greatest number of those in need. Yet, little research has been undertaken to investigate what strategies funders can adopt to incentivise more rigorous and transparent evaluation. Last, it is known that the focus of most third sector services is disproportionately influenced by single donors and major philanthropic institutions (e.g. the Bill and Melinda Gates Foundation). For example, 3% of the total number of UK charities account for 81% of the overall annual income of all charities [48]. However, there is limited understanding of how agendas in these larger organisations are created, and the subsequent influence they have on their larger network of stakeholders. Research is needed to understand the agenda-setting of philanthropic activities to enable an understanding of why certain social issues are addressed over others. To achieve this and to understand what types of giving is most effective, it was recently suggested that 'a science of philanthropy' is needed [49].

Further research on quasi-market provisions

Most research on ownership variation seems to focus on service *performance* or *quality*, with limited attention on what sector specific characteristics or constraints lead to the variation across different provider types. Therefore, a more detailed understanding of the *mechanisms* (for example through qualitative work) driving the observed differences across ownership is much needed. Further, past research in hospital settings has identified systematic variation among the types of

clients that self-select into for-profit, third, and public sector providers [50]. Unfortunately, the Care Inspectorate data did not allow us to test for variation among the type of users served by different providers, which should be explored in future research. Last, most research of this type is conducted in separate national contexts and on a limited range of service providers, thus making it difficult to understand how the wider policy and political contexts may influence ownership variation. For example, research in the UK and US [8, 51] tend to demonstrate quite clear differences in performance across ownership, whereas studies conducted within Scandinavia show more modest ownership variation [52, 53]. To understand how broader policy may be related to certain types of outcomes, comparative research is needed to determine whether and, if not, why, different kinds of variations can be observed in different national and regulation contexts.

Conclusions

The third sector is a central actor in modern society, and holds an important role in alleviating social issues related to, for example, the ageing population (Dilnot, 2017). However, the terminology and underlying meaning of 'sector' and/or ownership is subject to different conceptualisations across different fields of research, making it challenging to optimise regulation and funding to support the mixed social care provision effectively. The inconsistent use of terms relating to TSOs may facilitate ineffective use of the existing academic literature, as different fields use different terms for similar organisations. This thesis set out to improve the status-quo by undertaking focused, systematic, and rigorous work on the third sector delivery of social services. Focusing on research investigating the ability of TSOs to deliver social services, this thesis synthesised research within organisational sociology, public health, political science, and social policy to improve the understanding of sector in the context of social care services. Building on theoretical insights and a broad systematic synthesis of the literature, the thesis also

investigated sector variation among social care providers to find evidence suggesting that ownership may influence quality of care.

The findings of this thesis are significant for regulators and practitioners interested in the mixed provision of social care. First, the findings imply that funders and commissioners have the potential to improve existing incentives for TSOs to evaluate their services more reliably. Further, the findings suggest that the traits of TSOs, such as connecting with hard-to-reach-population groups, might be underutilised, in that TSOs often reported that the EBIs they were commissioned to implement did not match their target population. Last, the thesis identified systematic variation in quality-related outcomes across sectors, suggesting that ownership does matter in terms of social service delivery and outcome. However, more work is needed in this field, particularly in terms of identifying *mechanisms* that explain how sector-specific traits can be thought to influence social services. Understanding sector specific mechanisms have important implications for policy, in that an improved understanding of the distinct characteristics of different providers will enable policy-makers to address potential sector-specific challenges through regulation.

Bibliography

1. Kegeles SM, Rebchook G, Tebbetts S, Arnold E. Facilitators and barriers to effective scale-up of an evidence-based multilevel HIV prevention intervention. Implement Sci. 2015;10:50. doi:10.1186/s13012-015-0216-2.
2. The Social Care and Social Work Improvement Scotland (Requirements for Care Services) Regulations 2011. 2011. https://www.legislation.gov.uk/sdsi/2011/9780111012321/contents. Accessed 8 Oct 2018.
3. Bach-Mortensen AM, Montgomery P. What are the barriers and facilitators for third sector organisations (non-profits) to evaluate their services? A systematic review. Syst Rev. 2018;7:13. doi:10.1186/s13643-018-0681-1.
4. Bach-Mortensen AM, Lange BCL, Montgomery P. Barriers and facilitators to implementing evidence-based interventions among third sector organisations: a systematic review. Implement Sci. 2018;13:103. doi:10.1186/s13012-018-0789-7.
5. Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions. 2008. doi:10.1002/9780470712184.

6. Petticrew M, Roberts H. Evidence, hierarchies, and typologies: horses for courses. J Epidemiol Community Heal. 2003;57:527–9.
7. Bach-Mortensen AM, Montgomery P. Does sector matter for the quality of care services? A secondary analysis of social care services regulated by the Care Inspectorate in Scotland. BMJ Open. 2019;9:e022975. doi:10.1136/bmjopen-2018-022975.
8. Barron DN, West E. The quasi-market for adult residential care in the UK: Do for-profit, not-for-profit or public sector residential care and nursing homes provide better quality care? Soc Sci Med. 2017;179:137–46. doi:10.1016/j.socscimed.2017.02.037.
9. Cabinet Office. Public Services (Social Value) Act 2012. Queen's Printer of Acts of Parliament; 2012. http://www.legislation.gov.uk/ukpga/2012/3/enacted. Accessed 30 Aug 2017.
10. Cabinet Office. Social Value Act review. 2015.
11. Kelly J. Reforming public services in the UK: Bringing in the third sector. Public Adm. 2007;85:1003–22. doi:10.1111/j.1467-9299.2007.00680.x.
12. White GRT, Samson P, Rowland-Jones R, Thomas AJ. The implementation of a quality management system in the not-for-profit sector. TQM J. 2009;21:273–83. doi:10.1108/17542730910953040.
13. Barman E. What is the Bottom Line for Nonprofit Organizations? A History of Measurement in the British Voluntary Sector. Voluntas. 2007;18:101–15.
14. Maier F, Meyer M, Steinbereithner M. Nonprofit Organizations Becoming Business-Like: A Systematic Review. Nonprofit Volunt Sect Q. 2016;45:64–86. doi:10.1177/0899764014561796.
15. Dart R. Being "Business-Like" in a Nonprofit Organization: A Grounded and Inductive Typology. Nonprofit Volunt Sect Q. 2004;33:290–310.
16. Harlock J, Metcalf L. Measuring impact: prospects and challenges for third sector organisations. Volunt Sect Rev. 2016;7:101–8. doi:10.1332/204080516X14534734765005.
17. Harlock J. Impact measurement practice in the UK third sector: A review of emerging evidence. TSRC Working Paper 106 (Birmingham, University of Birmingham). 2013. http://epapers.bham.ac.uk/1800/1/WP106_Impact_measurement_practice_in_the_UK_third_sector_-_Harlock%2C_July_2013.pdf. Accessed 7 Apr 2016.
18. Pritchett L. It pays to be ignorant: A simple political economy of rigorous program evaluation. J Policy Reform. 2002;5:251–69. doi:10.1080/1384128032000096832.
19. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. PLoS Med. 2010;7:e1000217. doi:10.1371/journal.pmed.1000217.
20. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical Evidence of Bias : Dimensions of Methodological Quality Associated With Estimates of Treatment Effects in Controlled Trials. J Am Med Assoc. 1995;273:408–12.
21. Montgomery P, Grant S, Hopewell S, Macdonald G, Moher D, Michie S, et al. Protocol for CONSORT-SPI: an extension for social and psychological interventions. Implement Sci. 2013;8:99. doi:10.1186/1748-5908-8-99.
22. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. Br Med J. 2015;350:h1258. doi:10.1136/bmj.h1258.
23. McCord J. Cures That Harm: Unanticipated Outcomes of Crime Prevention Programs. Ann Am Acad Pol Soc Sci. 2003;587:16–30.
24. Bonell C, Jamal F, Melendez-Torres GJ. 'Dark logic': theorising the harmful consequences of

public health interventions. J Epidemiol Community Heal. 2015. doi:10.1136/jech.
25. Shelton RC, Cooper BR, Stirman SW. The Sustainability of Evidence-Based Interventions and Practices in Public Health and Health Care. Annu Rev Public Health. 2018;39:55–76. doi:10.1146/annurev-publhealth-040617-014731.
26. Shonkoff JP. Rethinking the Definition of Evidence-Based Interventions to Promote Early Childhood Development. Pediatrics. 2017;140:e20173136. doi:10.1542/peds.2017-3136.
27. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: The new Medical Research Council guidance. Int J Nurs Stud. 2013;50:587–92. doi:10.1016/j.ijnurstu.2012.09.010.
28. Durlak JA, DuPre EP. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. Am J Community Psychol. 2008;41:327–50.
29. Escoffery C, Lebow-Skelley E, Haardoerfer R, Boing E, Udelson H, Wood R, et al. A systematic review of adaptations of evidence-based public health interventions globally. Implement Sci. 2018;13:125. doi:10.1186/s13012-018-0815-9.
30. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review--a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy. 2005;10:21–34. doi:10.1258/1355819054308530.
31. Shepperd S, Lewin S, Straus S, Clarke M, Eccles MP, Fitzpatrick R, et al. Can we systematically review studies that evaluate complex interventions? PLoS Med. 2009;6:e1000086. doi:10.1371/journal.pmed.1000086.
32. Hawe A. T., P Shiell R. Complex interventions: how "out of control" can a randomised controlled trial be? Bmj. 2004;328 June:1561–3.
33. Department of Health. Liberating the NHS? 2011. doi:10.1136/adc.2010.205294.
34. Cabinet Office. Open Public Services: White Paper. 2011. http://www.cabinetoffice.gov.uk/sites/default/files/resources/open-public-services-white-paper.pdf%5Cnhttp://www.openpublicservices.cabinetoffice.gov.uk/.
35. Brennan D, Cass B, Himmelweit S, Szebehely M. The marketisation of care: Rationales and consequences in Nordic and liberal care regimes. J Eur Soc Policy. 2012;22:377–91. doi:10.1177/0958928712449772.
36. Krachler N, Greer I. When does marketisation lead to privatisation? Profit-making in English health services after the 2012 Health and Social Care Act. Soc Sci Med. 2015;124:215–23. doi:10.1016/j.socscimed.2014.11.045.
37. Dickinson H, Miller R. GP commissioning: implications for the third sector. Volunt Sect Rev. 2011;2:265–73. doi:10.1332/204080511X583913.
38. Billis D. Towards a theory of hybrid organizations Book section Hybrid Organizations and the Third Sector Chapter 3 Towards a theory of hybrid organizations. Billis Palgrave Macmillan. 2010;:46–69. http://eprints.lse.ac.uk/55260/. Accessed 26 Jan 2017.
39. Pestoff V. Hybridity, Coproduction, and Third Sector Social Services in Europe. Am Behav Sci. 2014;58:1–13. doi:10.1177/0002764214534670.
40. Alcock P. A strategic unity: defining the third sector in the UK. Volunt Sect Rev. 2010;1:5–24.
41. Anderson BB, Dees JG. Sector-bending: Blurring the lines between nonprofit and for-profit. 2017. https://centers.fuqua.duke.edu/case/wp-content/uploads/sites/7/2015/02/Article_Dees_SectorBending_2003.pdf. Accessed 30 Oct

2018.
42. Billis D, Glennester H. Human Services and the Voluntary Sector: Towards a Theory of Comparative Advantage. J Soc Policy. 1998. https://www.cambridge.org/core/terms. Accessed 21 Nov 2018.
43. Moore MH. Managing for Value: Organizational Strategy in for-Profit, Nonprofit, and Governmental Organizations. Nonprofit Volunt Sect Q. 2000;29:183–204. doi:10.1177/089976400773746391.
44. Rose-ackerman S. Altruism , Nonprofits , and Economic Theory. 1996;34:701–28.
45. Blomqvist P. The choice revolution: Privatization of Swedish welfare services in the 1990s. Soc Policy Adm. 2004;38:139–55. doi:10.1111/j.1467-9515.2004.00382.x.
46. Amirkhanyan A, Kim HJ, Lambright KT. Does the Public Sector Outperform the Nonprofit and For-Profit Sectors? Evidence from a National Panel Study on Nursing Home Quality and Access. J Policy Anal Manag. 2008;27:326–53. doi:10.1002/pam.20327.
47. Dilnot A. The burden of triumph: meeting health and social care needs. Lancet (London, England). 2017;390:1630–1. doi:10.1016/S0140-6736(17)31938-4.
48. NCVO. NCVO UK Civil Society Almanac. 2019. https://data.ncvo.org.uk/. Accessed 5 Jan 2017.
49. Fiennes C. We need a science of philanthropy. Nature. 2017;546:187–187. doi:10.1038/546187a.
50. Holom GH, Alexandersen N, Goldhaber-Fiebert JD, Hagen TP. Which patients receive surgery in for- profit and non-profit hospitals in a universal health system? An explorative register-based study in Norway. BMJ Open. 2018;8. doi:10.1136/bmjopen-2017-019780.
51. Karen E, Shen Y-C, Lau J, Schmid CH, Chan J. Hospital ownership and quality of care: what explains the different results in the literature? Health Econ. 2009;18 January:S37–54.
52. Stolt R, Blomqvist P, Winblad U. Privatization of social services: Quality differences in Swedish elderly care. Soc Sci Med. 2011;72:560–7. doi:10.1016/j.socscimed.2010.11.012.
53. Hjelmar U, Bhatti Y, Petersen OH, Rostgaard T, Vrangbæk K. Public/private ownership and quality of care: Evidence from Danish nursing homes. Soc Sci Med. 2018;216:41–9. doi:10.1016/j.socscimed.2018.09.029.

Appendices[6]

[6] All appendices are also available in the online versions of the publications.

Appendix A (Chapter 3)

Appendix A1: Adapted checklist for quality appraisal of cross-sectional findings

Items		Yes	No	Not applicable	Unclear
Introduction					
1	Were the aims/objectives of the study clear?				
Methods					
2	Was the study design appropriate for the stated aim(s)?				
3	Was the sample size justified?				
4	Was the target/reference population clearly defined? (Is it clear who the research was about?)				
5	Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?				
6	Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?				
7	Were measures undertaken to address and categorise non-responders?				
8	Were the risk factor and outcome variables measured appropriate to the aims of the study? (e.g. could you see the questionnaire?)				
9	Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously?				
10	Is it clear what was used to determined statistical significance and/or precision estimates? (eg, p values, CIs)				
11	Were the methods (including statistical methods) sufficiently described to enable them to be repeated?				
Results					
12	Were the basic data adequately described?				
13	Does the response rate raise concerns about non-response bias?				
14	If appropriate, was information about non-responders described?				
15	Were the results internally consistent?				
16	Were the results for the analyses described in the methods, presented?				

Discussion					
17	Were the authors' discussions and conclusions justified by the results?				
18	Were the limitations of the study discussed?				
Other					
19	Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?				
20	Was ethical approval or consent of participants attained?				

Quality (circle): **High** **Medium** **Low**

Reviewer: _____

Date: _____

Study: _____

Rationale and comment:
_____
_____
_____
_____
_____
_____

Appendix A2: Adapted CASP and JBI checklist

Checklist questions	Yes	No	Not applicable	Unclear
Was there a clear statement of the aims				
Is a qualitative methodology appropriate?				
Was the research design appropriate to address the research aims?				
Was the recruitment strategy/sampling appropriate to the aims of the research?				
Was the data collected in a way that addressed the research issue?				
Has the relationship between researcher and participant been adequately considered?				
Have ethical issues been taken into consideration?				
Was the data analysis sufficiently rigorous?				
Is there a clear statement of findings?				
Do the conclusions drawn in the research report flow form the analysis or interpretation of the data?				

Quality (circle): **High** **Medium** **Low**

Reviewer: _____

Date: _____

Study: _____

Rationale and comment:

_____

_____

Appendix A3: Top 3 cited barriers excluding the six studies of low quality.

Top 3 cited barriers
Lack of expertise and internal capacity (14/24)
Mismatch between funders requirements and appropriate goals (13/24)
Lack of financial resources (10/24)

Appendix A4: Themes excluding low quality studies.

Categories	Times cited
Factors associated with technical capacity and evaluation skills	**35-8=27**
Lack of expertise and internal capacity	17-3=14
Problems with data-collection and analysis	10-3=7
Inability to utilise existing data	3-1=2
Difficulty conceptualising and designing evaluation	1
Technical challenges	4-1=3
Factors associated with resources	**35-9=26**
Financial resources	14-4=10
Staff resources	6-2=4
Lack of time	6-1=5
Lack of resources to hire external evaluators	3-1=2
Resources (unspecified)	5-1=4
Staff turnover	1
Challenges in utilising evaluation systems and identifying outcome indicators	**25-4=21**
Difficulty developing and using evaluation tools	5
Lack of integrated systems to collect and analyse data	9-1=8
Challenges in identifying accepted outcome and impact indicators	12-3=9
Factors associated with funder requirements	**23-3=20**
Lack of funders requirements and support	3
Mismatch between funders requirements and appropriate goals	15-2=13
Poor proportioning of size of charities and funders requirements	1
Micro-management by donors	1
Funding insecurity (funding circles incentivising focusing on immediate outputs rather than long terms outcomes)	3-1=2
Factors associated with organisational culture and management	**24-4=20**
No perceived benefit to conduct evaluation	1
Staff resistance to evaluation	10-1=9
Perceived compromise between evaluation and service delivery	2
Lack of evaluation strategies and planning	3-1=2

Lack of feedback between board and management staff	1
Prioritisation	2-1=1
Lack of support from board and leadership	4-1=3
Evaluation not part of everyday practice	1
Other	**5-1=4**
Confidentiality of data	3-1=2
Lack of cooperation with stakeholders	2

Factors operating as facilitators

Factors related to receiving support to evaluate	**19-1=18**
Partnering with evaluation experts	3
Partnering with organisations working with similar activities	1
Technology availability and literacy to collect and analyse data	4
Benchmark data availability	2
Training of staff and evaluation literacy	5-1
Workable evaluation tools	2
Having sufficient resources to evaluate	2
Factors related to organisational culture and management	**18-1=17**
Understanding internal processes	3-1
To embed evaluation as part of everyday practice	1
Improve allocation of resources	1
Support from board and leadership	4
Have in-house evaluation staff	3
Having clear goals and evaluation strategies	2
Staff support	3
Positive perception of evaluation	
Factors related to the motivation to be accountable	**17**
Involvement of stakeholders to identify outcome indicators and evaluation goals	5
The motivation to influence policy	1
Compare work and outcomes to others doing similar work	1
The motivation to inform the sector as whole	1
Improve targeting of beneficiaries	1
Identify new approaches	1
Ensuring control and legitimacy of activities to stakeholders	3
The motivation to demonstrate and improve effectiveness of services	4
Factors around funder requirements and regulations	**5**
Funder requirements	4
Regulation requirements	1
Factors around economic sustainability	**5**
Using evaluation to be eligible for funding opportunities	3
Using evaluation as marketing	2

Appendix A5: Search strategy for included databases

Search string using ProQuest

Set#	Searched for	Databases	Results
S1	TI(charity OR charities OR "charitable organisation*" OR "voluntary sector" OR "voluntary organisations" OR "third sector" OR third-sector OR "community organisation*" OR community-based OR "community based" OR non-profit* OR "not for profit" OR "nonprofit" OR non-profit "social enterprise*" OR TSO* OR CBO* OR VCO* OR VCS* OR "civic sector" OR "civil sector" OR "social sector" OR CSO* OR "Non Government*" OR NGO* OR NPO* OR "civil society")	ABI/INFORM Global, Applied Social Sciences Index & Abstracts (ASSIA), International Bibliography of the Social Sciences (IBSS), MEDLINE®, PAIS Index, Policy File Index, Social Services Abstracts, Worldwide Political Science Abstracts	75454°
S2	MESH(charity OR charities OR "charitable organisation*" OR "voluntary sector" OR "voluntary organisations" OR "third sector" OR third-sector OR "community organisation*" OR community-based OR "community based" OR non-profit* OR "not for profit" OR "nonprofit" OR non-profit "social enterprise*" OR TSO* OR CBO* OR VCO* OR VCS* OR "civic sector" OR "civil sector" OR "social sector" OR CSO* OR "Non Government*" OR NGO* OR NPO* OR "civil society")	ABI/INFORM Global, Applied Social Sciences Index & Abstracts (ASSIA), International Bibliography of the Social Sciences (IBSS), MEDLINE®, PAIS Index, Policy File Index, Social Services Abstracts, Worldwide Political Science Abstracts	9888°
S3	ALL((barrier* OR facilitat* OR challenge* OR benefit* OR success* OR constrain* OR difficult* OR enhanc* OR	ABI/INFORM Global, Applied Social Sciences Index & Abstracts (ASSIA), International Bibliography of the Social	848811°

Set#	Searched for	Databases	Results
	influen* OR interfer* OR motivat* OR obstruct* OR problem* OR promot* OR restrain* OR restrict* OR disincentive* OR factor* OR capacity) NEAR/4 (evaluate OR evaluation OR evaluating OR assess* OR impact OR measurement OR measuring OR experience* OR learning OR performance OR “EBP” OR “evidence based practice” OR “evidence-based practice” OR outcome))	Sciences (IBSS), MEDLINE®, PAIS Index, Policy File Index, Social Services Abstracts, Worldwide Political Science Abstracts	
S4	ALL((barrier* OR facilitat* OR challenge* OR benefit* OR success* OR constrain* OR difficult* OR enhanc* OR influen* OR interfer* OR motivat* OR obstruct* OR problem* OR promot* OR restrain* OR restrict* OR disincentive* OR factor* OR capacity) NEAR/4 (evaluate OR evaluation OR evaluating OR assess* OR impact OR measurement OR measuring OR experience* OR learning OR performance OR “EBP” OR “evidence based practice” OR “evidence-based practice” OR outcome)) AND (MESH(charity OR charities OR “charitable organisation*” OR “voluntary sector” OR “voluntary organisations” OR “third sector” OR third-sector OR “community organisation*” OR community-based OR “community based” OR non-profit* OR “not for profit” OR “nonprofit” OR non-profit “social enterprise*” OR TSO* OR CBO* OR VCO* OR VCS* OR "civic sector" OR "civil sector" OR "social sector" OR CSO* OR “Non Government*” OR NGO* OR NPO* OR “civil society”) OR TI(charity OR charities OR “charitable organisation*” OR “voluntary sector” OR “voluntary organisations” OR “third sector” OR third-sector OR “community organisation*” OR community-based OR “community based” OR non-profit* OR “not for profit” OR “nonprofit” OR non-profit “social enterprise*” OR	ABI/INFORM Global, Applied Social Sciences Index & Abstracts (ASSIA), International Bibliography of the Social Sciences (IBSS), MEDLINE®, PAIS Index, Policy File Index, Social Services Abstracts, Worldwide Political Science Abstracts	3372°

Set#	Searched for	Databases	Results
	TSO* OR CBO* OR VCO* OR VCS* OR "civic sector" OR "civil sector" OR "social sector" OR CSO* OR “Non Government*” OR NGO* OR NPO* OR “civil society”))		

Search string for SCOPUS

Search 1:

TITLE-ABS-KEY (charity OR charities OR "charitable organisation*" OR "voluntary sector" OR "voluntary organisations" OR "third sector" OR third-sector OR "community organisation*" OR community-based OR "community based" OR non-profit* OR "not for profit" OR "nonprofit" OR non-profit "social enterprise*" OR tso* OR cbo* OR vco* OR vcs* OR "civic sector" OR "civil sector" OR "social sector" OR cso* OR "Non Government*" OR ngo* OR npo* OR "civil society")

Hits: 4713

Search 2:

TITLE-ABS-KEY ((barrier* OR facilitat* OR challenge* OR benefit* OR success* OR constrain* OR difficult* OR enhanc* OR influen* OR interfer* OR motivat* OR obstruct* OR problem* OR promot* OR restrain* OR restrict* OR disincentive*) AND (evaluate OR evaluation OR evaluating OR assess* OR impact OR measurement OR measuring OR experience* OR learning OR performance OR "EBP" OR "evidence based practice" OR "evidence-based practice" OR outcome))

Hits: 7,447,900

Search 3:

1+2

Hits: 1695

Search 4:

(TITLE-ABS-KEY ((barrier* OR facilitat* OR challenge* OR benefit* OR success* OR constrain* OR difficult* OR enhanc* OR influen* OR interfer* OR motivat* OR obstruct* OR problem* OR promot* OR restrain* OR restrict* OR disincentive*) AND (evaluate OR evaluation OR evaluating OR assess* OR impact OR measurement OR measuring OR experience* OR learning OR performance OR "EBP" OR "evidence based practice" OR "evidence-based practice" OR outcome))) AND ((TITLE (charity OR charities OR "charitable organisation*" OR "voluntary sector" OR "voluntary organisations" OR "third

sector" OR third-sector OR "community organisation*" OR community-based OR "community based" OR non-profit* OR "not for profit" OR "nonprofit" OR non-profit "social enterprise*" OR tso* OR cbo* OR vco* OR vcs* OR "civic sector" OR "civil sector" OR "social sector" OR cso* OR "Non Government*" OR ngo* OR npo* OR "civil society")))
Final Hits: 34

Appendix A6: PRISMA diagram

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-7
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	7
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	7-8
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	8-9
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7-9
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Additional file 2
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	8-10
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	9-10
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	9
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	9-10, Appendix

			A and B in Additional file 1
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7-10
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	9-10
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	15-16
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	15-16
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10-11
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Additional file 3
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Additional file 3
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	13-16
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	NA
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	15-16
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	15-16, appendix B and C in additional file 1
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	16-18

Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	19-21
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	16-22
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	24

Appendix B (Chapter 4)

Appendix B1: Prisma Diagram

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-6
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	7
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	3
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	8-9
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7-8
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Additional file 2
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	8-9
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	9-10
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	9
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	10

Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	NA
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	9-10
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	10
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	10
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10-11, Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Additional file 6
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Additional file 3 & 4
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	NA
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Table 1-3
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Additional file 3 & 4
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Table 4-6
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	16-20
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	20-21
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	22
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	23

Appendix B2: Search strategy

Search strategy via ProQuest	
Databases: ABI/INFORM Global, Applied Social Sciences Index & Abstracts (ASSIA), International Bibliography of the Social Sciences (IBSS), MEDLINE®, PAIS Index, Policy File Index, Social Services Abstracts, Sociological Abstracts, Worldwide Political Science Abstracts	
#1	ALL("evidence based practice*" OR "evidence-based practice*" OR "evidence-based intervention*" OR "evidence based intervention*" OR EBI* OR EBP* OR "evidence- based program*" OR "evidence based program* OR EBM OR "evidence-based service* OR "evidence based service*" OR evidence-based OR "evidence based")
#2	ALL(experience* OR attitude* OR perception* OR learning OR barrier* OR facilitat* OR challenge* OR benefit* OR success* OR constrain* OR difficult* OR enhanc* OR influen* OR interfer* OR motivat* OR obstruct* OR problem* OR promot* OR restrain* OR restrict* OR disincentive* OR factor* OR capacity OR enabler*)
#3	ALL(implement* OR adapt* OR adopt* OR adher* OR deliver* OR becom*)
#4	ALL(charity OR charities OR "charitable organisation*" OR "charitable organization*" OR "voluntary sector" OR "voluntary organisation*" OR "voluntary organization*" OR "third sector" OR third-sector OR "community organisation*" OR "community organization*" OR "community-based organisation*" OR "community-based organization*" OR non-profit* OR "not for profit" OR nonprofit* OR non-profit* OR "social enterprise*" OR TSO* OR CBO* OR VCO* OR VCS* OR "civic sector" OR "civil sector " OR "social sector" OR CSO* OR "non government*" OR NGO* OR NPO* OR "civil society" OR community-based OR "community based")
#5	1 AND 2 AND 3 AND 4

Appendix B3: Extracted quotes on barriers

Allicock et al. 2012

No quotes on barriers were presented in this article.

Amodeo et al. 2011

Certification Process

- "burdensome and time consuming"
- "high certification requirements,"
- "certification was difficult—constant fidelity checking"
- "having to record sessions creates rapport problems with clients."

Funding

- "not enough funding to implement fully"

Client Resistance/Client Non-Engagement/Lack of Suitability

- "client resistance to participation"
- "lack of participation from caregivers"
- "It takes a lot of follow up to get . . .adolescents to stay engaged, and sometimes it also means more follow up with the parents as well. . .so it is a challenge for those who are really not married to the program"
- "concrete thinking—CBT more appropriate for people who have insightful thinking"
- "clients could not stay focused because of outside relationships with relatives/friends"
- "in CBT groups, clients who miss the first week of the program when CBT is explained and goals are set."

Training

- "some staff without enough training are confused about the model"
- and [with so much staff turnover], "a barrier is keeping staff trained,"
- "it is easy for staff to "slip" from the MI framework; it requires constant clinical supervision; and a lot of time is required for appropriate clinical supervision."

Philosophical Differences

- "There is a perceived conflict between MI and the 12-step approach"
- "the transition to MI from the previous theoretical orientation is difficult for some clinicians due to different backgrounds"

Rigidity of the Model

- "little room for diversity in procedures"
- "very few cultural adaptations"
- "the developers had a very difficult time seeing beyond this one particular model, which was very specific. . .so there's absolutely no room to bring other orientations in"
- "clinician doesn't have flexibility for home visits."

Lack of Concrete Services

- "for homeless population, lack of housing is a barrier"
- "for co-occurring psychiatric and physical disorders, very limited resources"
- "lack of resources in county (no bus system, few jobs)"
- "project had limited access to substance abuse treatment."

August et al. 2006

No quotes on barriers were presented in this article.

Belza et al. 2014

Competing programming

- "For group classes we have dance, water aerobics, step aerobics, spinning, and the range of movement class from [another exercise program]. We have additional programs that are available at a cost, and those include our nutritional services; the EF classes; swimming lessons; different sports programs, and then small group training types of classes" (Age 30,11 years with the Y).
- "We acknowledge that space is an issue… They [wellness directors] see it as oh we already have [another program], our program for active older adults. Why would we want to do this one?" (Age 38, 8 years with the Y).
- "We [offer EF] off-site. We are not in our own building anymore. It was to save on rent…The big room is often taken up with children's camps and things like that" (Age 60, 7years with Y-affiliated site).
- "we no longer had the luxury of having two senior programs running because of space limitations."
- "we were having a time and space crunch. It wasn't anything wrong with the program *perse* but we're not going to take away our already very strong programs and try something new."

Limited resources and expertise

- "The staff did not see the benefit or the value to their people" (Age 45, 10years).
- "And getting our health and wellness directors to understand and not condemn it, like "What's in it for us?" (Age 38,8 years with the Y).
- "Where are we going to put it; who is the instructor going to be; who's going to pay for this, or where are the funds coming from" (Age 51, 16 years with the Y).
- "I know the whole issue is that people don't have time. There is a lack of staff. We have it here, too, and so I know some of the issues (Age 65, 6 years with Y-affiliated site).

Collins et al. 2010

No quotes on barriers were presented in this article.

D'Ippolito et al. 2015

Characteristics of Training

- "Had an emphasis on training from book learning rather than experiential learning"
- "There has been no formal training yet and there are inconsistencies in training as not everyone was trained the same way"
- "Training was not that good."
- "Some who go through training do not follow through, thinking it is too intense for them"
- "Training was five days and required travel"
- "Training requirements seemed unreasonable. Too much of staff's time."

Amount and Timing of Training

- "Peers weren't trained to work on the services provided through this contract, so the model wasn't really implemented"
- "Shelter staff weren't originally trained in MI."
- "Two years into operating with clients, training with MET and actual skill practice are very difficult. Skill building, even training once a month is not enough"
- "Additional trainings would have been worth it—not enough training."
- "Lack of continual education on the model"
- "Ongoing consistent training and the time to be able to do that versus I got my MSW and therefore I am a CBT [practitioner]."

- "Not having enough time at the beginning of the grant to train because we have to hit the ground running"
- "Lack of training at the beginning."
- "The amount of training that is needed to develop comfort with the model."

Agency Capacity for Training
- "Finding the time to get staff trained because it takes them away from their work."
- "Not enough money in grant for training or hiring master level clinicians."
- "The overnight and residential staff were hard to pull away to trainings."

Access to and Availability of Training
- "External trainings were far and few between, so when staff wanted training it was not always available."
- "Costs are outrageous for trainings and certifications, $45,000."
- "Getting to LA for the training so it could be taught to fidelity."

Training Related to Specific Staff Needs
- "A quarter of the staff did not have enough continuing education and struggled with the new model"
- "Not enough people were trained. We need agency wide trainings"
- "The case managers couldn't go to trainings, only the therapists."
- "Providing excessive trainings due to turnover"
- "Keeping staff up to date with what [EBP] means."
- "There are different education levels for our staff. It is diverse, so difficult to train all at once. They have different skill sets."

Demby et al. 2014

No quotes on barriers were presented in this article.

Dolcini et al. 2012

Assess
- Organizational capacity
    - "So it was more a matter of elimination as opposed to "YES, we know how to do this stuff." Because, in fact, it represented a totally new way of doing things for us and for the agencies we're working with, and it probably wasn't the best match. (ED)"
    - "I don't want to be so crass as if it was just a process of elimination, but in a way we looked at all the things they [CDC] were going to fund, and we said, OK, what can we not do, and crossed those off right away. Then there were a few that were left . . . And, then we had a dialogue in relation to our organizational capacity was as far as what made sense, and we ended up with [intervention]. (ED)"
    - "We were part of a [name] demonstration project. And, that was coming to an end, so we were actively looking to how to not lay people off. (ED)"
    - "And it seemed like our only option was to do it in this format, which was through, applying through the DEBIs. . . . Because previously we had been doing much more of a traditional outreach model. And, not done a structured program. (ED)"

Select
- Attending all trainings (getting everyone up to speed)
    - "And it was a while before the training, not everybody went to trainings right away, so it took a series of months for everybody at all the sites to actually go through all the trainings. And only once you go through trainings did we understand, did they really understand what this intervention was, and what we were gonna be expected to do. (ED)"

Prepare

- Staff turnover
    - "Well, there wasn't a lot of structural changes…. It was more hiring new staff…I think for us it was the challenge of starting up a new intervention. (ED)"
    - "We were in the process of implementation, within the first two months of finally getting our feet wet and starting our groups, one of the individuals resigned and moved…. so we were kind of hindered a bit…The next outreach worker also left our agency…It took us a long time to get a replacement. (ED)"

Implementation

- Clash between research and reality
    - "And so we feel very limited because … this is the enrollment session for the program, and so there's just this clash of, it's like a culture clash to me around the culture of research and the culture of the street. And the reality of these people's lives. So, it just doesn't seem like a very successful way to begin that relationship with them. (ED)"
- New to Staff
    - "Um, it was kind of a process of trial and error of how, um, a lot of the, a lot of pieces of the intervention were new to us and new to the staff. Especially like for the social…. What technique works best when introducing this to the client, on inviting people and friends in. It took us a while to kind of figure that one out and find something that was effective. (ED)"
- Monitoring and Evaluation
    - "We're evaluating in the sense of the basic three forms that we have. We have a pre and a post test and [a specific form that another agency requires]. Those are the three forms that I know of that are components of it [evaluation]…The concern, the question that we have is, How do you evaluate the data we provide? What does it mean? And, we don't have a system in place for that [evaluation]. (I)"
    - "…I know the evaluation form, but I sometimes don't believe in this form. Because the people don't say all… When, after the session, we talk to the clients, we ask them how they feel, what they think about the session. Most clients say that it's good, it's helpful. Most of the clients disclose their sexual orientation, their HIV status in the process. But this is not the reason [I think it is successful]. The reason is lower stress. Some people told us, "I feel good, no more stress." Maybe they don't or won't disclose their status, but they don't have stress anymore. And, it's good. This is the reason that this intervention is working. (I) (Although defining stress and improving coping skills are core elements of this DEBI, it is not a formal outcome for the intervention.)
    - "…I think there are probably [some behavior changes]. It's just a guess. I can't say that the intervention is effective. I can't say that it's not. (I)"

Feutz et al. 2013

No quotes on barriers were presented in this article.

Flores et al. 2016

Programs do not use their existing processes

- "The methodology is determined from the state agency . . . everything is determined by the agency, sometimes you must visit people, the hours are for training, the methodology to be used, instruments, all is defined by the agency."
- "We cannot recollect data. Database comes from the agency. All we ask users is to complete on the computer platform of the agency. The information is from the agency."

Lack of formal training in EBP – rely on literature

- "That's the beauty of this foundation. We work under the model of systems theory. Also we are reviewing ecological models. We try to be always updated on theoretical frameworks, as the situation of young people and children is changing through time. The phenomenon involved is not static, certain types of social vulnerabilities are increasing."

- "As you grow professionally you're using more complex frameworks. You start your bibliography from your university world searches, social workers look at the theory of systems, then all our frameworks come from there. Today it is more complex, because the agency is increasingly stronger in the development of conceptual frameworks."

Unable to appraise evidence

- "As the evidence is in Spanish, the language is important for decision making."
- "I think the biggest obstacle has to do with language, we are not able to read any language."

Accessibility of EBP information

- "That literature is easily accessible, with updated information."

Time constraints

- "I think the issue of lack of time is a barrier. The times in which we work are also barriers. The lack of time in which we organize the work does not give the possibility to read literature, attend seminars, and do more and more research."

Uninvolved in design

- "We are just executors, we implemented a project that is methodologically designed and we only indicate the timing and build the team that is in charge of the execution."
- "That is information that I do not drive, because all this comes from the agency, which has a department of study and they have seen international experiences. I do not know what other actors they have consulted."
- "The object of intervention in this program is defined by the agency, it is defined based on how this project will work. We have no interference in defining the target population. And this is defined by the agency. We only select the population based on the requirements defined by the agency"

Gandelman et al. 2012

Language

- "Well, the first change was to change the language, to translate the curriculum. . . . So all the curriculum had to be translated."
- "…one particular challenge that we're dealing with is the enrollment form. You have to ask them 15 questions which have six parts to each question roughly....So it's more of a clash between the research and the practical. We have a population that's um, hesitant to give out personal information, who if it seems anything formal or government driven or whatever, they're just very suspicious and unlikely to respond....our work…is based on building rapport and having trust,…relationship building is THE most important thing and right now we're coming in and having an interview, a formal research interview at the beginning."

Identity issues: sexual and ethnic identity

- "…class had to do with even the conversations they wanted to have…regarding relationships....The folks with the higher socioeconomic backgrounds were having different discussions around dating versus folks who were at the lower spectrum that might have been doing things for money...."
- "…if you have a 20-year-old sitting with a 40-yearold, it's very different, and we have class differences. We do have people who identify as gay or bi, and some who do not identify…and that's a challenge in having two different groups of individuals in the room. And we also have people who are not a person of color and would say they are…and they're really interested in the material and that's really hard."

Stigma

- "…because to be realistic, most of these clients [at our agency] do sex work for a living....But for them to accept that they're HIV positive, and the potential ramifications that it might have, legal

and economic....Because it's a small community.... so we have to open up a completely different day for them to feel safe....we have a more challenging time bringing them into this intervention because of that factor, their sex work. So, it's stigma about being HIV positive (and their ongoing sex work)."

Legal status and incarceration

- "…we pride ourselves on our efforts to keep confidentiality a priority, so it's a hurdle as far as advertising....it's kind of been a struggle that way, because we know they're out there but finding them in the best manner is what's hard."
- "They have to move around quite often. They have been pressured by the police. Some of them, specially the Latino population, they don't have documents to be in this country, so…[they]have to move and they are more persecuted than any other population. Also this population that you're seeing today, tomorrow you don't."

Drug use and homelessness

- "…this population is SO challenging…I don't really understand how other sites or how the original research group was able to get people to come in and come in again and come in again…the same group....it's like the most unstable population. These are people who, you see them in the street, you tell them about it, they say, yeah I'll come. You see them an hour before, they say yeah, I'll come, 15 min maybe you see them; we've actually…sent people out like the hour before to round people up, and they don't show up.

House, et al. 2017

No quotes on barriers were presented in this article.

Hunter et al. 2017

No quotes on barriers were presented in this article.

Honeycutt et al. 2012

Financial constraints

- "They thought that if we had some type of [party or family] event, [employees] would be saying, "Now you're asking me to take unemployment weeks, but . . . you're having an event." (Site coordinator)"

Kegeles et al. 2015

HIV prevention system factors

- The entire HIV prevention system affects intervention implementation
    - Funders
        - "When he [the coordinator] met with XXX, the guy from the health department…they don't want to pay for him to do activities that are associated with social events [a major Core Element of the intervention]…they only want to pay him to do the M-groups and the outreach."
- Knowledge about intervention
    - "The biggest difference between the way they do the model is that they have to follow the contract that the county lays out, which takes it far from the intervention…they have to do tons of things that are different from the model."
- Belief if efficacy of intervention
    - "They [the core group] don't plan anything as a group…the curriculum is ideal but unrealistic for real life as they know it…people are busy… so the coordinators do the events…she thinks that that part of the program, the core group planning stuff, is a joke and needs to be changed…maybe in a place where there is nothing to do but pick your nose, but in [a large city], it isn't going to happen."
- Desire to change agency's existing approach
    - "They had been trying to get it [MP] going…or something like it….they had a lot of people coming, but it wasn't focused…it didn't look like the model…the agency didn't

really care. They do outreach once a week at the clubs, and they do referrals for [HIV] testing and STD testing… they go through and give out safer sex kits to anyone who will take them from them....[but] they haven't redeveloped the kits for Mpowerment."

- Accountability for work
  - "When she [the supervisor] goes to talk to [the ED] and she tells him about stuff that is going on with the project, he just sits there and nods…he doesn't provide any direction for her and is more apt to err on the side of supporting the coordinators rather than questioning if they are a good fit or not…this makes it difficult for her…when she thinks that they may need to be fired."
- Appropriateness and capacity of individuals for coordinator positions
  - "She hired someone on recently who has a marketing background… who is tuned in with what is attractive and fun… but when it comes to promoting something…he is like 'I don't know, I don't know…I don't know'…she thinks that even with the training, and with the weekly meetings…he feels a bit overwhelmed…he is good at databases and stuff like that…but publicity, marketing, recruitment…it is very hard."
- Organizational stability
  - "He [the funder] thinks that the project needs to go to a different organization…the agency isn't stable enough."

Community factors

- Geography
  - "'Trying to compete with San Francisco [to attract YGBM] wasn't working" [respondent's CBO was located in a county only a few miles away]
- Sociopolitical context
  - "Publicity was hard because they couldn't even have a website that was geared toward gay men because the county is so conservative and the funder [the county health department] didn't want to risk creating a commotion."

Intervention factors

- Intervention characteristics
  - "When he went to Atlanta [to a conference], he saw [a presenter connected with the MP research] and he talked to him and went to his presentation, and at the time they didn't like the [intervention] because they thought that they [the researchers] had all this money to do it and it wasn't the real world…because they were talking to CBOs that were working on shoestring budgets and they don't have lots of money for training and planning like they [the researchers] had."
- Adaptation issues
  - "'The question of fidelity is something that they talked about a lot… the boxed [DEBI] interventions are great, but what people really need is more TA about how to effectively adapt these interventions while retaining the theoretical core. They [the agency] needs to build their capacity to understand the internal logic of the M-group piece of the intervention so that they can say 'here is the logic of this activity, and the behavior it is seeking to address…here is our target population for this intervention…how do we change M-groups for this target population while retaining fidelity to the original design?'"

Kimber et al. 2012

Having clear leadership

- "Flounder"
- "Run around in circles"
- "Need to request extended deadlines and support in completing tasks."

Supplying needed resources

- "Like when I was starting on that working group, I was also meeting [the demands of] our accreditation process. So, you know, I kind of had to find the balance between doing both of

those things, knowing that the one—the accreditation, had a specific timeline. So, I said to the group, you know, we are about to do this (accreditation), and if you want to go on to the next steps, it will have to wait until another time, because these other things are up."

Doing this differently

- "My issue is that some of the trainings are not even provided by our own staff. Yes, we need professional training… but using our own resources, like we have done for some of the Triple P training, and if we could do that for some of the other EBP trainings that will be rolled out, that would be good."
- "There are still people in the agency that are involved in this (CT) that shouldn't be, because they don't get it, and they don't try. It is clear that this isn't their priority; it's more just on their plate. Because there are some people that do this really well, and then there are others who struggle, and then the strugglers keep coming, saying I can't get it, I can't get it. Well, that's because you can't do it."

Lattimore et al. 2010

Program logistics

- Readability of material ("small fonts, literacy level above the recommended fifth- to sixth-grade reading level, or appropriateness of graphics, pictures, and examples")
    - "The example in the book or on the website doesn't include retired people; therefore, their situations are not similar. The examples don't look like me and they have different lifestyles…."

Program theory

- Tracking thoughts related to physical activity
    - "Participants are not completing the 'keeping track of thoughts' and 'keeping track of PA.' They [participants] all commented that they do think about PA all the time, but the tool to track it isn't simple or friendly."

Program philosophy

- Adhering to screening and eligibility requirements
    - "Moderate intensity is like you are rushing to get to the bus before it pull away."
    - "The functional level is so low for some participants that if they can consistently engage in activity over the course of the program, that is huge."
- Assessing light versus moderate activity
    - "Everyone has a different perspective of what light activity is versus moderate activity."

Lundgren et al. 2011

Lack of fit with current organizational practice

- "The other thing that we, we had to do is this is meant to be a closed group, in other words you have your women, they always come to the same group, it's always the same women. And that was not doable for us at all. So we had to adapt it…to be an open group, and again…one did not have the sense that they [the funder]were happy about any of this for us. So…I don't know that we have a whole lot of fidelity in the pure sense going on here, although I don't think we changed any of the curriculum or anything."

Lack of fit with organization's population

- "…in this program we don't do the full [EBP] because it's with [client population], and we had to tone it down a little bit, cause the true [EBP] is just too aggressive. It doesn't really work with them…."
- "You know we have diverse populations and so some of the – a couple – there's seven different types of ethnic populations that we're serving. Three of them I think the [EBP] fits very well. The other four populations, we have to do a lot of modifications. So that's why when you say – you're asking me the scales – it's actually a scale for all seven. But if for three of them we didn't

need to modify very much, maybe at a two or a three level, but the others are more a six or seven level."

- "We have over the past years, because…predominantly our population is [client population], and…so we've made adaptations to…how you move people through [EBP]. That looks different culturally."
- "We did…they'll tell ya…with the [EBP] about the four legs of the stool…and all the different components that need to be there, and [client population] is one that when we talked to the folks they said 'we're not even sure if this…can completely work with [client population] because of the lack of supports that they have in their lives.'"
- "It's really kind of ingrained in…'if that's not working then try something else.' I know we have done a lot of the case management while at the same time doing the[EBP], so it's not really [EBP] not standing alone, but…[specific client population] have so many different needs…that we try and support them when we can, and that's why we…add some of the case management alongside the [EBP]."

Lundgren et al. 2012

No quotes on barriers were presented in this article.

Marahaj 2010

Philosophical orientation

- Level of congruence
  - "Program Philosophy Score (3): There continues to be a basic IDDT orientation articulated by all sources interviewed, however elements of the model that remain problematic, e.g. lack of group DD treatment, provision of substance abuse treatment outside the agency, and generic client goals/plans, have not progressed significantly, despite their centrality to the EBP's core principles. Although consumer and family interviews had been scheduled, the interviews did not take place. An agency brochure that communicates the IDDT philosophy has not yet been developed. Some stakeholders from other systems who refer clients to the agency (and elsewhere) do not yet exhibit a good understanding of the model's core components. (LF1, 24M Fidelity Report) Program Philosophy, score was 2. In speaking with staff, at various levels of the organization, a focus appears on getting to the destination of IDDT. Helping staff see this program in a long-term, if not permanent, manner, will likely help LF3 to more fully implement IDDT. Identifying ways for the team to communicate this to one another and to reviewers will assist this organization with fidelity. (LF3, 12M Fidelity Report)"
- Norms, values, and beliefs
  - "Rather than outreach, motivational interventions, and other engagement techniques for people in the engagement and motivational stages of recovery, the team members seemed to talk more about setting boundaries for clients by not helping them find new housing when they lose old housing due to drug use. (LF2, Baseline, Fidelity Report)"
  - "They are pretty down on this implementation as well. They see it as a barrier to their own productivity. They see the training as long and useless, and they really don't think this program has been very helpful to them. They don't see clients in the stage of change the way that IDDT promotes. They don't really see much hope for many of these clients. (LF3, 18M Implementation Monitor Summary)"
  - "Some team members were uncomfortable with IDDT's support for attempts to reduce rather than immediately eliminate substance use."
  - "There are however, a few resistant team members, according to the program leader, who are not comfortable with letting clients make their own choice." This attitude was highlighted in the 6M consultant/trainer's interview for LF3: "Their attitudinal set going into this was very paternalistic and controlling," in addition to "they don't really value consumers input all that much," while "some staff talked about clients in a non-professional and negative way."
  - "That is the ongoing issue that this team has been dealing with since the beginning of this project. Trying to figure out where to put their IDDT team has been a huge issue

for this team. I believe I read this is the largest mental health service agency in the state, so perhaps it's that bureaucracy that is preventing this organization from finding the proper place for its IDDT team. (LF2, 12M Implementation Monitor Summary)"
    - We've talked all along in the implementation process that if you anchor the program in a person or persons it's not going to fly. You have to anchor in the culture and they're one site where it's not anchored in the culture (HF1, 6M Trainer Interview)

Leadership

- Role of leadership
    - "I think there is definitely a misconception from program leaders. They are a little out of touch with what their staff does on a daily basis."
    - "I think the program leader has not been extremely effective...he speaks very passively about he is just doing what his administrators tell him to do; he is not invested in the model. He doesn't care about the model. He doesn't particularly believe that the model will work better than what he is already doing and he disagrees with us in terms of how closely they adhere to the model (LF3, 6M Implementation Monitor Summary)"
    - ""The lack of clarity around this important part of the multi-disciplinary team is symptomatic of the way the program came to be structured."
    - "Agency administration was astute in recognizing the value of IDDT over parallel services but underestimated the complexity of the model and what it would take to implement it. One way that this underestimation played out was the assignment of IDDT program roles to staff who were either not "with the program" or not adequately skilled to carry out those roles. (LF1, 6M Implementation Monitor Summary)"
    - "She was just describing those behavioral acts as things they were motivational counseling. So, I don't really think she understands what it is. It is a very important part of stage wise interventions, it is a very important element of outreach and engagement but I just don't think she understands what it is so she is not able to lead her staff and tell them what motivational counseling is. (LF2, 6M Implementation Monitor Summary)"
    - "As if the Program Leader or Registered Nurse went to Borders, bought a book on "here's this great new thing to do with your group" and just start doing it unbeknownst to the rest of the agency and that is the way it will probably happen and if somebody said you can interview the upper administration a year from now and say "so how is the…work you're doing here?" and they can say "what's that?" (HF1, 6M Program Leader Interview)"
    - "About [the CEO], ….in spite of whatever he brings to the table personality wise, he got that agency out of the red and into the black and his fiscal management was a clear strength. If you got underneath the questions, they were fiscal questions and that over time the fiscal viability of the team has been demonstrated to him, he's now sort of with the program. …….when he gets up there endorsing it, he's not endorsing fidelity, he's endorsing the viability of it, the outcomes and the fiscal viability (HF1, 24M Trainer Interview)"
- External leadership
    - "Steering committee acts as steering committee for both the ACT and IDDT implementations. I suspect that having a steering committee overseeing both practices might be diluting the steering committee's mission too much."
    - "The former implementation monitor has noted that the consumer and family member involvement at this agency is very minimal. For example, she said that the agency picked a consumer and a family member who would be nonassertive for this steering committee."

Adaptability

- Consensus and buy-in
    - "I think another one is just kind of well, "this whole intensive thing is just a little bit much for me, I kind of like the idea of a 40-1 deal where it's more generic and I don't have somebody that wants to kill themselves every week," and so I think there are some

people who are hanging in on the team because they're good people and they're team players but they're not particularly invested in where they are. (LF1, 6M Program Leader Interview)"
    - "...So there is a range in attitude toward the model with some people being somewhat enthusiastic about it all the way down to people where you can just tell they are rolling their eyes and they don't want to be involved. It is just a real mix of how they can really get this model going. I asked a couple of people why they think they were on the team and they had no idea, they were just drafted basically. (LF2, 6M Implementation Monitor Summary)"
    - "The administrators at this agency, they've been more of a source of barriers I would say than strategies. They really are bitter; they don't feel like this is a project that is very well funded……..they would rather just have the State give them the money rather than"
    - "giving us money to train them. You can tell that they are skeptical about the effectiveness about the model, how much it costs, so there is definitely consensus building going on with the administrators right now. (LF3, 6M Implementation Monitor Summary)"
    - "It does not appear that the Implementation Resource Kit (IRK) materials have been used outside of intensive trainings with the consultant/trainer in ways that could address some of the clinical issues. My impression is that once the IRK trainings were over, the materials got shelved. (LF1, 6M Trainer Interview)"
- Lack of changes in agency practices
    - "This agency still thinks we know best and we are going to take care of you,"
    - "there is some resistance at this site particularly to the idea that there are things they need to change."
    - This is an organization that has clearly decided, "We're not gonna continue." I think their behavior around that, their chart notes, their efforts, their supervision, have been moving away from that. They are not attending to that. So, they, knowing that they're moving away from it, I think that they stopped putting an emphasis on it. (LF3, 24M Trainer Interview)"
    - The emphasis on the status quo was evident in the more traditional approach to services, and perceived as strengths by two low fidelity sites. As described by the program leader for LF3: "Well, there are areas of strengths, I guess, because that is the way we have always done business here." The 24-month fidelity report capture this 'old school philosophy' as reported for LF3: "This group is focused on use of a 12-step informed process. Little is shared about psychiatric symptoms or coping techniques. Abstinence is expected. The group has been meeting for many years."
- Resistance to change from practitioners
    - "Program leader stated that the use of stages in notes was not being encouraged because that "jargon" was not used throughout LF3."
    - "They're still doing the same old treatment that they've always done, and they're not trying to incorporate some of the new stuff into there."
- Lack of clear role delineation
    - "Who's driving the bus is a question that really needs to be answered,"
    - "the role which should have embodied both IDDT expertise and championship was seriously unfilled."
    - "The consultant/trainer has been aware all along of the mistakes that the agency leadership has made in the direction that the implementation has taken. He tried to steer them toward a more advantageous route but they had their own ideas. (LF1, 6M Implementation Monitor Summary)"
- Lack of time
    - "It was difficult to integrate the practice into daily work because of high caseloads (around 40) and lack of time to perform IDDT activities that are not reimbursed….and the case manager responsible for the most IDDT clients declined IDDT mini-team involvement b/c he did not feel he had time for an extra meeting every week (not

surprising considering the high caseloads). (LF2, 18M Implementation Monitor Summary)."
    - "The program leader has received training, and she helps monitor the quality of the IDDT implementation. She also arranges for supervision and training. However, she is not totally empowered because of conflicts with trainer and lack of time. She does not service IDDT clients herself, and she does not have time to working to improve IDDT at LF2. (LF2, 18M Implementation Monitor Summary)"
    - "They just simply don't have the time because of their large caseloads,"
    - "Productivity remains paramount and time for training or other learning is restricted."
    - "staff has a lot of things on their plate outside of this project."
- Lack of funding
    - "Comprehensive services- this is an area where deep cuts in funding have affected the program's ability to reach high fidelity, particularly around residential and vocational services….For instance, where there are resources in the community/state, such as the IMR CCOE, the agency has been unable to mobilize the energy and time to access them. This is a product of downsizing, budget crunches, etc. that place a burden on the agency and makes it difficult for staff to tackle new projects. (LF1, 6M Implementation Monitor Summary)"
    - "In order for me to be able to do something, no matter what a great intervention technique, I have to be able to make money. If I'm not making money, then I can't do it no matter how wonderful it is and it would almost have to be designated state by state, because Medicaid and how it funds is different in every state you go to. ..It all boils down to we want to do good clinical treatment, evidence based treatment but at the same time, that's why we are having a problem. (LF3, 6M Program Leader Interview)
    - "LF2 administration has made it clear through their indecision about how to form the team that they do not have the money (or maybe the will) to truly implement IDDT right now. Finally, DMHA is affecting the implementation in that it is not offering the same incentives for IDDT that it does for ACT. And….voc rehab provides a disincentive by not allowing supported employment reimbursement if the IDDT client can't be kept employed long enough. (LF2, 18M Implementation Monitor Summary)"
    - "The themes in the areas where we have low fidelity to the model usually relate to cost and revenue production."
    - "Consultant/trainer had misgivings about the fiscal policy/philosophy of the agency administration beginning with the first interview with the CEO…..thinks that fiscal uncertainty at HF3 probably lead to turnover of team personnel and threatened stability/sustaining of the program despite its clinical soundness. (HF3, 24M Final Implementation Report)"

Knowledge

- Lack of training
    - "If you haven't been through the training, it's like speaking a foreign language to other clinicians. I think that they could easily pick up on it, but it's not something they use on a regular basis. It just doesn't make the work process easier to use in-terms that the person you're talking to doesn't understand. (LF3, 18M Program Leader Interview)"
    - "It is hard to squeeze it all in. There is a lack of time. They have productivity concerns that are a barrier for them to receive more training. I think they would like to learn these things, and they would like to make time, but they also do not want to lose agency incentives on their productivity. They don't want to be punished for receiving training in the model. (LF3, 12M Implementation Monitor Summary)"
- Lack of mastery
    - "Although clinicians verbalize the principles of client choice, the lack of individual/personal goals actually incorporated into the treatment planning and documentation belies insufficient mastery of the principle."
    - "Staff seemed to be trying to use motivational interventions in some notes, but they were still not getting it."

    - "Team staff reported that they tailored their treatments to clients' unique needs and goals. However, this individualized treatment was not documented in the progress notes of the charts we reviewed. In fact, some of the charts' progress notes contained identically worded typewritten statements with blank spaces where each client's name was written by hand. (LF2, 12M Fidelity Report)"
- Lack of clinical sophistication
    - "It speaks some to the lack of sophistication that exists not only on their team but on so many of these case management teams around the state."
    - "These are very young case managers and are maybe in their first position and really do not have an idea about what to expect."
    - "There was minimal understanding of these areas in previous fidelity reports. She also said she wasn't sure that the team leader herself was grasping motivational interviewing or stage wise interventions."
    - "This group, not only had no background in substance abuse, they had no background in anything, like 3 of their 5 case managers it was like their first job out of school, they had nothing so they needed as much mental health background and substance abuse. (LF1, 6M Trainer Interview)"
    - "Team members continue to display a range of sophistication and expertise with regard to these techniques, commensurate with their background and experience, although the level of understanding apparent in their description of interaction with clients has increased. (HF1, 12M Fidelity Report)"

Lack of priority for supervision

- "The team's case managers receive supervision on a more informal basis,"
- "No clinicians receive structured, weekly supervision from a practitioner experienced in IDDT that is client-centered and explicitly addresses the application of IDDT to specific client situations."
- "Supervision score =2. According to the team staff, no formal supervision time is scheduled, although staff is free to seek out supervision as needed. The recommendation is that IDDT is a model that requires a great deal of clinical skills, so all IDDT team staff could benefit from structured weekly supervision. (LF2, 24M Fidelity Report)"

Martinez et al. 2014

No quotes on barriers were presented in this article.

Murray et al. 2014

No quotes on barriers were presented in this article.

Owczarzak et al. 2011

Intervention level

- Lack of fit
    - "Lack of fit"
    - "After having conversations with [the prevention director] about it, the people we work with are so random. There are some people we see all the time, and some we don't see for months, and then some people move" (Female, Associate Director, 9 years with agency)."
- Unique characteristics of target population
    - "Well, when you go into the jail system you can't leave any condoms. You can't leave literature. Initially the staff people [at the jail] wanted to sit in on a group, and that kept everybody not being open and honest. But then, they [staff people—JO] stopped coming. You can tell those women don't want to say what's going on with their life. They may be doing something right up in there [in jail], and they don't want to say. And the lack of their ability to go outside. Some women that are in those institutions whereby they're court mandated to be there, and they can't have passes and stuff. So it's really

hard to gauge a behavioral change if they really aren't doing anything. [Female, direct service staff,10 years with agency]"

- Time intensive
    - "We talked about it [the possibility of implementing a DEBI—JO] and talked about it, and tried to do it and just couldn't. It was— basically we discovered that it was full time. It was a full time job just to do that program. [Another staff member] was really, really interested in trying that and doing that but 2 years later we just were like it's, it's, because those, those staff out there were one man. They do everything. They do needle exchange groups, bar outreach, internet outreach . . . [S]o if we had had a specific person that could be dedicated just to doing that program that would've been just great but [not]with the limitations we have. [Female, Associate Director, 9 years with agency]"
- Lack of desire to shift priorities
    - "I think one of the challenges, and this was a decision I had to make this week, actually, on the budget. DEBIs, because of their sustained pieces, it's really important to have the pieces like the food and the pieces like a $40 gift card. The reality is, it's very expensive to run them. And if you have a limited budget and you need to pay your people, you need to have pieces of your overhead in there, your indirect costs in there, all of a sudden then you're making choices. I had to make a choice this week: do we want to run the program or do I want to have tape, or do I want to have a travel budget so that I can have staff development for my staff? I chose we need to run the program for next year. [Female, Director, 1 year with agency]"
- Intervention conflicts with organizational identity
    - "It was too focused. It's too rigid of a focus when you are segregating down to that [young MSM—JO]. I thought back to our group and I [realized that] I've got more women, young women that come to group than the young men. And if we are facilitating this, if we are going to be doing this and this is going to be beneficial, it needs to be broader or it needs to be more encompassing. It needs to be a space—regardless of how you identify or what you do or how you do it—it's going to be a place for you to go. I expressed that it was just too rigid, too rigid to sit down and have a group of 10 young women that identifies as lesbians and focus this thing directly at them because you get diversity within that group. [Male, Youth Specialist, 8 years with agency]"

Organizational level

- Staff turnover
    - "I know a challenge was staffing. Early on, we'd get somebody hired, trained and the momentum going and then [he] leaves. And then hire and trained and momentum going and now left again. This is like the third staff person in 2 years. So that's been a huge challenge. And I think a lot of that has to do with probably the credentials of the person that we're looking for in the position. Ideally, we like somebody that has at least started the process with AODA [alcohol and other drug abuse—JO] counseling, just the group facilitation piece. And then, the salary maybe is the issue—that we're not paying an AODA counselor's salary. Then they're getting a job in an AODA field and then moving on. And then also too, our grant target has one of the primary population's Hispanics. So getting somebody bilingual also has been a huge challenge. Bilingual individuals, especially AODA counselors, are pretty hard to keep because they're sought after. [Male, Prevention Director, 9 years with agency]
- Make decisions regarding resource allocation, use of staff time, and agency-level priorities
    - "I do know that prevention dollars for education were cut and therefore programs . . . When I was a prevention education specialist, that program was cut. It was the HRH, High Risk Heterosexual funding we were getting. That was cut, so there was not that much money for any more education. And that was probably close to the time that maybe they just decided not to implement [VOICES/VOCES], I don't know. Because initially, the agency has to take in the cost of whatever these programs are going to cost before the CDC or someone else gives money for it. [Female, Prevention Director, 5 years with agency]"

- Issues with funding/funders
    - "The demographics around [the agency] have been shifting. . . .We were hired to do prevention in the Latino and Latina community, and there was no mention of the fact that there still needed to be gay outreach"

Program level

- DEBI adaptation, modification, and fidelity
    - "I don't think anybody thinks that adapting is a bad reason, is a bad idea. I think that—this is just my impression—that the state thinks that this is something that the CDC developed [so] it's their responsibility to come up with a Latino one . . . the state isn't going to take it on . . . For an individual agency to take it on it would be a huge amount of work. [Female, Prevention Director, 6 years with agency]"
    - "We looked at MPowerment really closely. [My colleague] and I wanted to do it, but knew we couldn't do it like it was written— to the extent that they wanted it done . . . [We] thought maybe we could adapt it, but we were told we shouldn't adapt it. [Female, Associate Director, 9 years with agency]"
    - "P: Some of these other things that they suggested don't seem 100% vital to it and they don't fit for us anyway. So we are just going to change those but we felt like we really kept true to the core of the initiative, like the sessions and have it be group led, and building community. That was most important that we kept that. I: So you decided on your own what you saw as the core elements of this program? P: Yeah we did it and we did not go back to the people and say this is how we are going to do it. Partially because we knew they might say, "Well, that is not going to work." So . . . we knew in our heads probably what they were going to say but we didn't really want to hear them say because we knew we had to do this way anyway. We are going to have to do [it] this way and we just did it. [Male, Associate Director, 1 year with agency]"
- Lack of follow-up and accountability built into the DEBI program.
    - "P: There wasn't a ton of follow up. I think that a lot of the assistance was one sided, us reaching out to them. There wasn't necessarily a lot—or any, that I can remember—follow up on their end. Like, "Alright you came to this thing. What have you guys done?" The door was open but just us to talk to them. I: Do you think that that would've been helpful if they would've made more effort to contact you? P: Yeah, I think it would have been helpful. I think that it definitely would've helped with accountability if we didn't do anything and they call us and we may feel guilty like, "Oh my gosh! We totally went to this training nine months ago and we haven't done anything with it." Also, it could just help us reinforce. . . we may have then asked other questions that didn't feel important enough for us to call and ask this question. [Male, Associate Director of Prevention, 1 year with agency]"
- No opportunity to discuss implementation process
    - "I think it will be good if the same people who were involved in facilitating the training came [back for] meetings and discussed certain things [such as] what else can be done differently. I know that's a lot of money—those strangers came from Colorado, but I think giving us continual education about the curriculum and how people implemented it . . . the experiences of other people doing it. That would be helpful. [Male, prevention specialist, 8 years with agency]"
- Training facilitators
    - "I think there was a lot of book knowledge . . . It sounded beautiful, but the reality was, "OK, this is going to be hard." For MPowerment, it was like someone coming from California, where there is a population that we are trying to reach, and that's a big minority number. Being MSM themselves, it was more effective [than] having someone having the book knowledge and just talking about it, saying, "This is what you do next" and "This is what you will do at the end." And I'm thinking, "OK, it sounded nice." And you get out there thinking, "I can do this," but then you sit down and see, "OK, where are we going to get these people from." [Male, direct service staff, 10 years with agency]"

- "'[The trainers] were very focused on, well, 'This is how you spend your 40 hours. You have plenty of time to get this and this done.' But they didn't take into consideration that we don't have that luxury. We are one person and have to do the testing, the driving back and forth, and the office stuff" (Male, prevention specialist, 5 years with agency)."

Owczarzak 2012

Provider perspectives on the need for EBP

- Problems with DEBI program
  - "I think it was sort of more about style than about content. I think the DEBIs are good. I think everybody needs to understand that evidence-based practice needs to be happening. You need to say why you are doing something, and you need to say what the something does to the folks that you have given it to. But the way in which it was rolled, at least the way it was perceived . . . at the time: I had co-workers and friends in the community who were at other CBOs and their perception was that their programs were devalued because they weren't evidence based . . . And it really felt like a top down edict to a lot of people. [Female, 40s, director level, 1 year with agency]"

What works in HIV prevention

- Tension between evidence-base and their own knowledge
  - "[I]t just didn't look feasible to me and I'm by no means an expert, please, but I kind of had a gut feeling for what would work and what wouldn't. And that [DEBI] I was never a fan of. The people that presented to us – please, no offense – were too white to me. They were too – they didn't seem to know my community and you are coming to tell me this is good for my community and you don't know my community? How do you know it's good for my community? Tell me how you came to that conclusion. So I was not – I had a lot of questions and concerns and comments. [Female, 40s, direct service provision, 14 years with agency]"
  - "For the Mpowerment [a community-level intervention for young men who have sex with men – JO], it was like someone [trainers – JO] coming from California, where there . . . [is a] population that we are trying to reach. [But, there] it's a big minority number. And being MSM themselves it was more effective [than] someone having the book knowledge and just talking about it and saying this is what you do next, and this is what you will do at the end, and I'm thinking, 'Ok. It sounded nice'. And you get out of there thinking, 'I can do this', but then you sit down and see, 'OK, where are we going to get these people from?' And it's hard to get them to come in once. Three times the same people? [Female, 30s, direct service provision, 10 years with agency]"
- Realities of implementation
  - "I think a lot of [DEBIs], in my experience with them, kind of look good on paper and maybe work well in kind of a research setting, but then transferring it into an agency that's doing the work on the streets, sometimes it just doesn't transfer over very well . . . I don't think the populations we serve can be as predictable as some of the DEBI's kind of make them out to be. There's just a huge gray area of how people respond or will respond to an intervention. [Male, 50s, director level, 9 years with agency]"
- Lack of resources for evaluation
  - "I don't necessarily have any data or anything, but from where I sit in my office, I can see the HIV program's door. So I see youth going in there. And not by name, obviously, but [my coworker] who does HIV tests will just kind of like anecdotally be like, 'Yeah, so I saw 10 youth in this last 2 weeks'. [Female, 20s, direct service provision, 1 year with agency]"
  - "I think it's very effective because you understand what you are supposed to do, and it wasn't that difficult to start it. It was something that we had the population, and everybody wants to get together and do a social thing, even coming into the office or meeting somewhere. I think it felt like it was interesting for them because it was a time for them to meet and one-on-one, to talk about the difficulties that each of them face

and to talk about risk reduction. [Female, 30s, direct service provision, 10 years with agency; emphasis added]"

- Lack of knowledge regarding evidence of effectiveness
    - "I don't how you would measure the effectiveness . . . I guess it depends on how you look at what area of effectiveness. I mean, this program has a couple of different things that you can get from it. If you're learning something from the program at all, any thing about your risks, it's effective. That's one measurement for me. Another measurable outcome for me is if you're actually able to work toward steps toward changing your behavior, working towards any step, whether it be your first step, your second step, or your final step. That's effective to me. Again, it focuses on baby steps. Another measurable outcome is when you can achieve a goal, whatever it is that [you] decide to achieve . . . How you overcame your own barrier . . . and brought change about in your life . . . If you have changed your thinking process in general on just how you cognitively deal with specific things . . . We've seen them – we've seen people on every stage. So I think just by them being here, it's effective. [Female, 30s, director level, 5 years with agency]"

Payan et al. 2017

Planning

- Personnel capacity
    - Provider inertia
        - "I think sometimes older providers have the way they treat it, and sometimes they are stuck. They do it their way versus new ways or learning how others do it."
- Professional development capacity
    - Lack of knowledge or experience with revised clinical guidelines
        - "It's outdated now because we have new guidelines. I need to look at this to confirm. When we started last year, it was JNC 7 and recently, there were some changes."
    - Lack of time to train personnel
        - "There is room for improvement in the clinics and certainly for increased provider education or documentation."
    - Lack of consensus around medication practices
        - "There isn't consensus around medication. If there's guidance on that, that would be helpful."

Implementation

- Personnel capacity
    - Management, provider, and staff retention issues
        - "We personally have had issues because of transition with leadership so it's been hard to implement programs when there's a constant flux in leadership so that has been one of our personal barriers."
    - Competing responsibilities
        - "These projects are great, my own personal thing is I am frustrated because I wish I had more time . . .I wish I [had been] involved in the initial planning so we could have implemented sooner. I just don't think I'm doing the project service with my time restriction"
- Technological capacity
    - Variable costs and delays
        - "I think that because of tremendous unforeseen problems with our EMR system, there have been long delays in the implementation of the hypertension specific decision support measures and that both of those factors have made it hard to expand the project to the volume of patients that would have made it more meaningful during the months of the study period."

Pemberton et al. 2012

Implementation

- Need for greater leadership support
    - "Administration needs to learn more about MIP, what it entails, and the work conducted by staff. Lack of knowledge results in limits and less support. The organization's leadership doesn't place a lot of importance to MIP…there should be more orientation in this area [with] the Directors and Administration. [They] should be knowledgeable about MIP program goals and be able to understand the humanistic work that is being done."
- Culture
    - "Even though my staff has been working with IDUs for over 8-10 years and they have expertise in terms of sitting down with a client and getting him to work through his risk-reduction goals. For the client--the main importance … is the drug and getting the drug—that's a major barrier. I guess the other agency will tell you the same thing"
    - "They [MIP clients] will disappear but they will always come back. They may not be as structured as other people because of the drugs, but they always come back to us. MIP is client centered so you cannot make them do anything; you work with them little by little and motivate them to do better."
    - "Activities/referrals do not always go as planned. MIP is time-consuming because of who we are working with[IDUs]. The flow from one session to the next does not always happen as planned."
- Capacity
    - Human resource constraints
        - Inadequate or limited program staffing levels
            - "[Agency] requires more staff to run the program as intended--maybe one more staff because I need at least two outreach staff on the streets...Then I need a Counselor and Case Manager to work with MIP because that is one of the core elements of the intervention."
            - "Although case conferences are conducted and there is some mentoring with writing clinical notes, "staging" clients, etc., staff definitely need more supervision"
            - "[Agency] model is different from what is recommended in the MIP curriculum. Each staff member does both counseling and case management—using the recommended MIP model would ease the burden."
        - Staff skill deficits
            - "I am still convinced that MIP is better implemented by clinical staff. For example, even in how you respond to clients, there is a clinical response. I believe MIP would be more effective if implemented by people with a stronger clinical background"
            - "[I was] "learning as I go." I would go out with my manual (MIP Implementation Manual) wherever I went and would be flipping through the pages. Am I doing this correctly? Have I covered everything? Then I gained my confidence. I still have that thing in my trunk…."
- Content
    - Meeting program fidelity (time constraints, quality, levels of service)
        - "[MIP] is a tough intervention to implement because each component can break down into multiple sessions. It is good, but it is long. When you work with IDUs you need to have an intervention that is more direct-- short --and can address all those components in a less amount of time."
        - "We have an intervention that has 7 sessions and every session can be broken into 5-6 meetings--other sessions. Then you have a client who has a big substance abuse problem and he may not comply with the meetings that have

been established. So basically…we can work with a client over 6 month period [as per MIP guidance]. During that time, let's say the client was attending nicely and complying nicely for three months and something happens and then he disappears and the work that we have done in 3-4 months…and we are close to the end... it just disassembles in our face—so that's a major barrier."
        - "Meeting target number of clients as indicated in the program contract is a challenge. Some clients take 10 minutes to get through a session, and others take 3 days"
    - Implementing MIP as recommended
        - "When [original researchers] performed the study, they had a lot of case managers and staff to implement it. Now the reality of any agency working with HIV is that they don't have a big staff; it is a few staff with various responsibilities…that is the reality in HIV[programs]—maximize the staff. When you work with an intervention as structured as this, you have to take the client to another level to really complete it…..we want to really be effective; we want to create an impact, so you have to do it right, and when you do things right it takes a little bit more time. So that's the thing."
        - "MIP is challenging to implement structurally. All sessions are not conducted with all clients. This is because [the] client does not follow a schedule or session flow"
        - "MIP is challenging in terms of doing it as is…..Client attendance is not reliable. If clients are no shows for over a month, you have to go back again and start from the beginning."
        - "[We] considered the intervention presented as opposed to how it is in practice. For example, the intervention says 7 sessions but each session can take multiple contacts and run for as long as 6 months"
- Community
    - Funder specification on client eligibility for MIP
        - "The fact that qualifying MIP participants must have shared within the last 12 months has been a huge challenge. MIP staff has had to turn away individuals because of this clause."
    - Inadequate community support/structural systems to support MIP clients
        - "So far the great challenge we have encountered is the lack of mental health programs that will take clients. It is very hard to link them for psychiatric evaluations."
        - "Until basic needs are met client will not commit to recovery. We do not directly work on substance abuse but work with their top issues….housing, food, urgent care, etc. That is beauty of MIP."
        - "While the community is not directly involved, they are not opposed to the MIP program either. [Agency] has a commitment to the community and is open to ideas to facilitate the effectiveness of the MIP program."
        - "The community does not buy in as much…community has more important priorities on their plates than MIP. Also, the addiction community is transient."
        - "[The] community sees population in a very negative way; so there is not much community support"
- Multiple categories
    - Inability to provide clients with all needed services
        - "Sometimes clients come because they have a lot of case management needs—housing, food supplies, employment—things that need to be attended quickly…and that's part of the challenge because sometimes you cannot really resolve as you wish—because you are calling other agencies and they do not have an opening for the client—then the client may feel frustrated and he may feel that you are not really doing enough for them. Of course you try to explain to them....we always make it clear to the client that we don't promise anything. What we can promise is that we will work with them and we are going to try our

best. We never tell the clients that we are going to give them this and that because that is not a good thing. When it comes to mental health it is challenging as well to find some referrals for clients with mental health [needs]."
        - "[Agency] could improve on relationships with other agencies. Agreements are done but there is no "warm connection" with other agencies. MOA's are on paper but the relationships are not always there or work out as intended. Sometimes we go to agencies that we know that we have agreements with and we do not have anything in common."
    - Inability to provide clients with monetary incentives
        - "would like to see incentives used for certain unobligated things. For instance, money to supplement costs for receiving certain services where there is a small fee. For example, if you [provide] a referral for a client to seek medical care or a medical examination but they do not have the $20.00 co-pay, then they don't go—[I]would like to see incentives used like that."
        - "If clients do not have identification in a lot of places they cannot receive services—it is like they don't exist—clients have to have a picture id....especially since 9-11. Clients need monetary help to get id's."
- MIP Program Implementation Support
    - Lack of helpfulness of funder
        - "MIP TOF was provided but no additional trainings or refreshers have been offered by the funder. We are working with what we know."
        - "MIP TOF training, materials and documents provided by the [funder] was helpful. Since then no refresher training or new strategies have been offered to MIP staff."
        - "Funder is not trained in intervention- "Technically I don't know they can assist because of their own lack of knowledge. [Funder] relies on us— the implementing agency."
        - "Funder} lacks knowledge and understanding of the intervention and
        - therefore cannot provide guidance on implementation of MIP."

Adaptation

- Lack of knowledge on how to adapt EBIs
    - "The message for everyone trained has been that there are certain elements of the intervention that cannot be changed. You must follow the curriculum otherwise the efficacy/validity is compromised."
    - "Depends on the core elements of the intervention. Use core elements as your rules on what is allowed with adaptation."
    - "The core items need to be in place and not changed. There needs to be flexibility, but the essence must remain or else it becomes a totally new program."
- No guidance on adaptation
    - "A lot of what I learned in the DEBI workshops is that the core of the program cannot be adapted—that there can be variations.. but certain things that cannot be changed."
    - "[The] guidance is general with a focus on keeping core elements. It is not specific to populations. We are always told "don't touch the core elements."
    - "Some guidance on adaptation was received through trainings, but not specific to MIP"
    - "[I was] told we can adapt but not "how" to adapt."
    - "No guidance was received on adaptation. All modifications—the terminology and forms—were done by agency staff without guidance"

Petrescu-Prahova et al. 2016

Implementation themes

- Lack of organizational support
    - "[The residential site] did most of the recruiting…there was no encouragement by the staff there for people to continue in the program." Female instructor, 3 months experience

    - "…they will not allow instructors to substitute for other instructors at different Ys, which I find kind of silly…" Male instructor, 1 year experience
- Weaknesses of training
    - "I don't think that [EF instructors] have enough knowledge after a two-day training to take on the responsibility of people's health." Female instructor, 2.5 years experience
- Fidelity issues
    - "I think that the fitness checks are a good idea, but it is hard to get the time to do them…" Female instructor, 2.5 years experience
    - "It's kind of a challenge because I can't do two things at once. Because my classes are only half an hour apart, I have to kind of do the fitness checks either before class or during class and have my sub[stitute] lead the class while I'm doing them…" Female instructor, 1 year experience
- Cost of program delivery
    - "When we came into the grant, we actually had a phenomenal opportunity to really impact those…[who] were very low income…" Male master trainer, 2.5 years experience
    - "…when we first started EnhanceFitness, I think that it was kind of hard to get the higher-ups in the [non- Y] company to buy into it. There were no grant dollars involved to get it up and running." Female staff, 5.5 years experience
    - "I worked with [name of funder] to get this program in our [non-Y] facility, and so people didn't have to pay… The grant ran out, and when we moved to this location there was a charge." Female staff, 7 years experience

Maintenance themes

- Organizational structure for program delivery
    - Poor staff
        - "We went through several really, really bad instructors that the students were complaining [about]… they were going to quit because the ability of the instructors was so poor." Female instructor, 5 years experience
    - Scheduling
        - "That was the hardest part for us scheduling-wise, because we only had so much space and so many hours that we can offer it." Female staff, 7 years experience
- EF tasks as part of employee responsibilities
    - "There is so much paperwork that it is just ridiculous." Female staff, 2 years experience
    - "…as long as you keep up with [the paperwork]…and turn it in regularly, it's not a problem." Female staff, 5.5 years experience
    - "Well, EnhanceFitness is the only one that requires paperwork. We don't really do paperwork for any other class." Female instructor, 1.5 years experience
- Cost for class participants
    - "For some people who are on fixed incomes, the cost of EnhanceFitness can be difficult… I think if there were more opportunities and it was cheaper for people—or it was covered by insurance like [name of program] is— it might get more people involved." Female instructor, 1 year experience
    - "Some of my people … complained about the fact that they had to pay to take this class." Female staff, 7 years experience
    - "I think that the Y does a fantastic job by not charging. If you're not a member, it is a suggested donation." Female instructor, 1 year experience

Pinto et al. 2015

Costs

- Staff time and precollaboration
    - "Nothing is taken off their plates"
    - "This collaboration was added on top of their initial goals and tasks"

Balancing the needs of the CBO and research

- Balance during collaboration

- "Multiple divisions and levels require a great deal of coordination and for the CBO staff they have to add this study to the pile of other tasks."
- "dealing with the needs and concerns of both parties is where the challenge lies"

Communication

- Communication and future collaboration
  - "one challenge was that we weren't able to have face time together. All communication was over email and never meeting in person is a little tough and not ideal."

Ramanadhan et al. 2012

Resource constraints

- Staff capacity
  - "It's a lot of paperwork! . . . First they said you had to do an evidence-based program. And then they said, 'Okay, but their evidence-based evaluation isn't enough, on top of that, you have to do ours.' . I've always found it too restrictive . Participant 5 (Focus Group, Boston)"
- Material resources
  - "The evaluation . tends to be more process evaluation, than outcome evaluation . Some grants are fairly rudimentary in their outcome, in their evaluation. They just look at, 'Did you provide the service?' and they don't really look at who was served or what the outcome of the person. You know, the people we serve, were they any better off at the end? Although, I think everybody would like to get to that point, nobody really knows exactly how to do it, given the limited resources to do it.—Participant 6 (Key Informant, Worcester)"

Challenges with program adaptation

- "I'm thinking like, across the country. So, if something may work really well in the Midwest, you have to think of inner-city populations of people .you're gonna definitely find, clearly a difference in what works, scientifically, here." —Participant 1 (Focus Group, Boston)
- We can adapt all we want, but then we have to really say, 'Is it really . are we maintaining and keeping what it's supposed to be?' Cause when you translate, you kind of lose it. Not the whole thing, but you lose pieces of it." —Participant 7 (Key Informant, Lawrence)
- We don't change the concept, we change sort of the way maybe it's delivered.—Participant 8" (Key Informant, Lawrence)
- "That's all, I find that's a challenge sometimes when . a funder kind of says . 'Stick to the letter!'"—Participant 1 (Focus Group, Boston)"

Organizational culture

- "Sometimes, if there's an important individual who's interested in evidence-based programs, or who has heard about something that works, they're more likely to do it. But I don't think it's part of the culture" .—Participant 6 (Key Informant, Worcester)
- I think a lot of people believe that solutions need to be developed locally. And, so, what I've seen more often is trying sometimes to reinvent the wheel too many times.—Participant 9 (Key Informant, Worcester)

Thomas et al. 2014

No quotes on barriers were presented in this article.

Vanderpool et al. 2011

Lack of fit

- "should be taken back where they came from. They don't fit X County"
- "I think the biggest challenge that I felt was that when I was looking through the list…some of these things looked very specialized and were really well designed. I really didn't feel that a lot of those specifically, not just regionally or geographically, were a good fit for our economic demographic of very low income families."

- "I'm also very aware of the literacy issue for many of my patients. Folks don't need all the details, they want to know how this works at my house, how can this work in my life, so I'm always looking at that angle for almost anything we do here."

Negative perceptions
- "expected"
- "overwhelming"
- "challenging"
- "restricting"

Volunteers are unfamiliar with EBIs
- "Most of our [church] members are 'average' citizens who are unfamiliar with the [EBIs] concept and want to use what they believe will work."

Difficult to adapt
- "I had to think of cramming 4 weeks of 'Give Me Five' programming into 45 minutes worth of time to get the one month of programming in."
- "I felt a little distraught throughout the [adaptation] process because I didn't think I was doing it right."
- "I felt like it was a lot of extra work that I didn't know if it was going to pay off in the end."
- "…it was adapted so much to the point where it would probably be a fine line to say whether or not you could even relate the impact we had on our families to any of those particular interventions."

Recruitment
- 'We had an "If you build it, they will come mentality" which didn't work'.

Funding
- "…it [the intervention] will never be a large scale success unless the money and timeframe are more realistic to more adequately show the full impact the program had on the community.'

Lack of information on EBI use in a similar context
- "We're looking at evidence-based models now [for physical activity in the schools] …I think it gives you a definite argument as to why a program should be run, because you have the research to back you up, but again it's hard because some of those programs you do have to tailor. There are not very much of those programs tested and run in Appalachian West Virginia. I think we have a different ball-game here…We're all dealing with the same diseases across the US, but there is a definite culture difference here and sometimes it's very hard to breakthrough that barrier. You have to get very creative."

Veniegas et al. 2009

Preimplementation phase
- Acquiring information
  - "I didn't realize how strict the people are to give out the information. . . . You can't just buy it, you have to go through the training. . . . We had missed like already two trainings that were happening. . . so the next time around was going to be too late."
- Assessing fit
  - "Evidence-based interventions don't fit with the agency because of the complexity of the target population . . . homeless, Latino and African American, men and women . . . because of the issues that they deal with, such as language
  - barriers and culturally related issues, some of these interventions really don't happen to apply specifically to this population"
- Preparing organization and staff

    - "We know that prevention is different for [HIV] positives and [HIV] negatives. Now you are asking us to not only adapt and tailor but to develop another curriculum to complement this."
- Securing technical assistance for intervention selection
    - "didn't know how to do it either. There were no other resources available."

Implementation

- Securing technical assistance for implementation
    - "without having any evidence that that's gonna make it any more or any less effective."
    - "We've tried to combine sessions because of retention, and we've gone through a couple of series where we've combined a couple of sessions together, and the clients have made it through it"
- Conducting process evaluation
    - "We're spending all our time trying to learn what to do, then adjust what we're gonna do, and by the time we're actually doing it, then our staff member leaves."
    - "The big barrier with retention involving adults 24, 25 and up is people work and have lives. People go on vacation. They'll show up to one session, maybe two, but to show up for all three is really hard unless we're providing . . . a big incentive"

Maintenance and evolution phase

- Supporting staff for continued implementation
    - "I don't think there has been training out there that gives the staff ability to do that."

Appendix B4: Extracted quotes on facilitators

Allicock et al. 2012

No quotes on facilitators were presented in this article

Amodeo et al. 2011

No quotes on facilitators were presented in this article

August et al. 2006

No quotes on facilitators were presented in this article

Belza et al. 2014

Match with the Y mission

- "Adopting an evidence-based program fit well with the Y goals and standards and was congruent with the Y mission. The Y mission includes providing programing that helps to improve the health and fitness of older adults" (Age 30, 11 years with the Y).
- "It's back to the spirit of mind and body of what the YMCA does (Age 33, 11 years with the Y)."
- "We knew it was an evidence-based program, one that fit well with the Y goals and standards" (Age 30, 11 years with the Y).
- I think it falls into our focus areas i.e., healthy living for seniors and social responsibility as well. We're being responsible when we provide those types of exercise programs" (Age 29, 5 years with the Y).
- "Our strategic initiatives and our strategies roadmap for our association states very clearly that we will have a growing focus on expanding our senior membership and increasing our programing to meet the needs of the aging population"

Organizational support

- "The association office asked us if we would be interested. Because of the clientele we have, we have lots of seniors, of course I stepped up to the plate and said: "Yes, definitely let's try this for our seniors." That's how we got involved" (Age 41, 14 years with the Y).
- "I think having a focus and support from junior management is important" (Age 38, 8 years with the Y).
- "I think you need to have an Executive Director or CEO really understanding what it means to deliver evidence-based programs" (Age 38, 8 years with the Y).
- "I think it's always good thing to bring something new in. It was driven by the Y of the USA. And then also I was asked to do this by our Health and Wellness director of the main branch" (Age 51, 15 years with the Y).

Match with the target population

- "the residents [in assisted living] were good candidates because they were at a point where they had not been exercising. We could start at the beginning and see where their progress was which would not have been if we brought it into our Y and tried to offer it to our regular seniors".
- "We had been looking for an older adult program because we have a large aging population in our community. It has been an age bracket that had been underserved at our Y" (Age 30, 11 years with the Y).
- "We are known for the fact that we offer programing that is valuable to the community and to the seniors in the community. EF is one of those programs" (Age 63, 2 years with Y-affiliated site).
- "We are in close proximity to (low income housing) and so this is a very easy place for them to come. It's convenient for them. If we're talking about people that are low income and don't have money for public transportation, it makes it very easy for them to do something to take care of their healthy living" (Age 63, 2 years with Y-affiliated site).
- "It is completely appropriate for many health seekers and people who struggle with becoming more active or staying active" (Age 38, 8 years with the Y).

Financial

- "When we started it, we started with the [state department of health]… they gave us a grant basically along with other YMCAs in [the state] with all of those being downstate. They basically paid for my staff's training and they sent us. I think that they also paid for all of our equipment. They were a huge, huge partner in this and for us being able to start EF when we did" (Age 33, 11 years with the Y).
- "The [state] contacted us and we've been working with them for some other programs. They offered to help with the initial training, and that's where we learned about the program" (Age 30, 11 years with the Y).
- "We offer financial assistance. Based upon income I can give participants a certain percentage off the price of the class. And then based upon some of the grants that we have been given, I can give them even a higher percentage off. We do the best that we can to really make it happen for them. I don't like saying no to anybody" (Age29, 5years with the Y).

Champions

- "[Champions] just do it on their own. Nobody asked them to do that (in reference to setting up fitness check areas). They just love the community that EF provides and obviously the physical benefits. They want to capture anybody that comes into class and really helps them feel that same way."
- "I am a go-getter and if I hear something, I go after it because it is beneficial to our residents" (Age 65, 6 years with the Y-affiliated site).
- "For me, personally, it was something else for me to offer to the seniors. I absolutely love working with the active older adults" (Age 41, 14 years with the Y).
- "The more and more I learned about it, the more I loved it. I didn't really know of any other like evidence-based programs for older adults. I really liked the pre and post-tests that they did. It just seemed like a great program" (Age 33, 11 years with the Y).
- "I told my boss about it and how I thought it would be beneficial. I told our members about it because I wanted to get them on board and get them excited. I did anything I could when the (grant sponsor) people came over. I did everything I could to promote our space" (Age 60, 7 years with Y-affiliated site).

Novelty

- "Back then [when EF was adopted] EF was kind of an experiment. There were only a few sites in the country offering it, I believe, and so I thought that it would be nice to be part of that group."
- "I thought it would be something different, you know? I thought it would be more different and something that we could offer to our seniors (Age 41,14 years with the Y).
- "It was just something new and exciting, evidencebased. It was everything we wanted (Age51, 11 years with the Y).
- "I think it's always good thing to bring something new in" (Age 51,16 years with theY).
- "I just wanted to have a varied program offering, and I thought this would fit…I wanted to keep the people who come here happy with our center. I want to give them a variety of things, and so I don't want anything to be stagnant" (Age 60, 7 years withY-affiliated site).

Invitation to partner

- "The opportunity to work with and partner with an outside agency to help address another portion of our population definitely interested me."
- "I think that I thought it looked like a great program. Our partnership at the [state department of health] was so strong. They really wanted to help the YMCAs start it" (Age 33, 11 years with the Y).
- "Sometimes the senior centers request us to do a program. That is kind of how it happened. It was just really good timing when we started EF because they were requesting that we come and do some different things. We thought it would be perfect and so it just kind of fell into place" (Age 33, 11 years with the Y).

- "I was working with a grant writer at [a university]. I was looking for something that we could get through a grant. This is the something she came up with" (Age 60, 7 years with Y-affiliated site).
- "Someone [YMCA staff member] called [my manager] and said: "We have this EF class and would you want to be our pilot program?" And she said "Absolutely, yes! That's how it all started" (Age 51, 11 years with the Y).

Program specific facilitators
- "…an easy sale as it was proven to improve things"
- "It is evidence-based and has got solid backing. It has a proven track record and can meet the needs that are out there" (Age 30, 11 years with the Y).
- "I won't touch anything that does not have data or an evidence-based curriculum, especially as related to chronic disease management" (Age 38, 8 years with the Y).
- "It is an incentive to bring people in when they know that you have a program that is known throughout the country. It's a recognizable name. You are branded already" (Age 63, 2 years with Y-affiliated site).

Collins et al. 2010

No quotes on facilitators were presented in this article

D'Ippolito et al. 2015

Characteristics of Training
- "Received trainings from the person who wrote IMR."
- "Training was excellent"
- "Good training which was thought out and planned well."
- "The shift from communal training to individual training based on skills and experience"
- "Trainings were helpful particularly because of the use of videos."
- "We had extensive training, a 3- or 4- part series on CBT"
- "Having an intensive external trainer."
- "Training outreach workers to reframe questions"
- "Cultural sensitivity training."
- "The curriculum itself is easy to understand, and could be implemented without formal training."
- "The people who do the training are very supportive."
- "This training spurred our staff to want other kinds of training."

Amount and Timing of Training
- "Training is ongoing which helps to learn new model and correct implementation"
- "Continued follow-up trainings with staff"
- "Trainings so frequent it kept it fresh for staff."
- "We go to trainings a lot."
- "Having training at the onset of the project."

Agency Capacity for Training
- "The ability of staff trainers to provide training."
- "We had the funds and resources to hire an expert to provide training."
- "The agency's willingness and commitment to continuously train staff."
- "Training in M.E. is offered for the entire agency, so it brings everyone on the same page."
- "Training partners are most helpful."

Access to and Availability of Training
- "The yearly training provided through SAMHSA. We come back energized and it reinforces that we are doing the right thing."
- "There are a lot of agencies around to provide free or low cost trainings."

- "Community relationships—we're able to get in right away for trainings."

Training Related to Specific Staff Needs
- "The clinicians were highly educated so they didn't need much training"
- "We had a clinical supervisor on staff who was trained in ACRA"
- "Staff are well trained and well educated."
- "My previous experience with the model through internal trainings."
- "New staff are required to have education/training series on IDDT."

Logistics
- "Staff went to UC San Diego conference for training."
- "Training on IDDT that the agency provided"
- "Bringing in DBT specific training."

Demby et al. 2014

No quotes of facilitators were presented in this article

Dolcini et al. 2012

Assess
- Fit of program
    - "It was a natural fit into the population we serve. And a large amount of our population are injection drug users. It was an easy fit. They're not that many DEBIs that apply to injection drug users, so there wasn't much of a choice. (ED)"

Prepare
- Having other services offered
    - "The reason that I think that we work well is that we have other supportive services in-house. And, I don't know if all the [other funded] agencies have other supportive services as well. If they're not, then it will be a big challenge. [Intervention name] is not a stand-alone program. You need to have other things in place. (I)"

Implementation
- Ability to adapt program
    - "Um, I think it's gone really well. I think the trick with the intervention is really being able to adapt it and to tailor it. I think if we stuck to the curriculum very strictly it would be tough. But that there is room, I think, to aim it towards, towards the population, and I, I really feel like [the clients] are getting stuff out of it. (I)"

Feutz et al. 2013

No quotes on facilitators were presented in this article.

Flores et al. 2016

Ability to adapt program
- "There is evidence that allows us to go feeding back program execution. That feedback should be interpreted according to the characteristics of each program. The technical bidding which gives the agency are only guidelines not rules, so we can improve our intervention. It is up to each executor."

Gandelman et al. 2012

No quotes on facilitators were presented in this article

House, et al. 2017

No quotes on facilitators were presented in this article

Hunter et al. 2017

No quotes on facilitators were presented in this article

Honeycutt et al. 2012

Leadership support

- "Pastor . . . was wonderful when it came to supporting the program. . . he was really there for me. (Site coordinator)"

Access to technical assistance

- "We'd call [TA provider] for a lot of things that we didn't understand, as far as the implementation of programs, . . just not taking it away from the evidence-based program. (Site coordinator)"

Kegeles et al. 2015

HIV prevention system factors

- Planning for intervention before implementation
    - "The room they've ended up using is quite sterile and housed within the AIDS project."
- Evaluation of intervention's functioning
    - "And now they are starting to think about who they aren't reaching…who are the people that you aren't reaching out to…he doesn't want it to be just about handing out flyers…he spends a lot of time getting the guys to reflect on what they are doing with the project."

Kimber et al. 2012

Clinical transformation project manager

- "Having the leadership of the project manager as sort of overseeing the whole process of clinical transformation was really effective. Having that presence early on in the group meetings, that was very effective in sort of setting the stage for the process and how the process works."

Having an inclusive change culture

- "For me, I think the most pertinent thing has been the front-line staffs' involvement in the clinical transformation process. So, having them in the working groups, having them involved in the implementation directly, and have a real say in terms of how the interventions have come in and been developed and incorporated into their work has been imperative and significant to the process."

Supplying needed resources

- "The management of the agency has been very supportive of our trying to manage CT as well as other commitments throughout the year."
- "It is so easy to have meetings conflict that are equally as important as other things. That hasn't happened because of the messaging and permission from the agency to make this (CT) a priority."

Lattimore et al. 2010

No quotes on facilitators were presented in this article

Lundgren et al. 2011

Ability to adapt program

- "Not intentionally if we did, it depends on each person. We don't believe in cookie-cutter help, so it depends on the situation. The need is the circumstance and the individuals. I mean…I'm not going to use [EBP] completely all the way, depending on the length of recovery that an individual has, depending on their attitudes, their self esteem levels, et cetera. We haven't changed it… it's evidenced based, for God's sake. You gotta use what has been proven to be successful. But you also have to modify it for [geographic location], for the urban or the rural. I

mean, so those are some of the differences I think that make it so that you have to modify it at all."

- "So we have…made some adaptations in just how we go about the treatment, and our openness, following mostly the same principles, but we have had to make some changes throughout the process."
- "We have over the past years …. made adaptations to the whole process and how you move people through [EBP]."

Lundgren et al. 2012

No quotes on facilitators were presented in this article

Marahaj 2010

Philosophical orientation

- Level of congruence
  - "There is a pervasive culture of respect for the client selfdirection at the agency that is very much in the spirit of recovery and consistent with the IDDT model."
  - "Focus on and respect for client choice continue to be organizing principles of treatment at the agency,"
  - "Printed materials reflect a philosophy consistent with the principles of integrated treatment, including a stagewise, client-centered approach to seamless and comprehensive SA and MI services."
  - "There's a unified philosophy that supports the principles of the IDDT EBP," while at HF3: "There is a strong foundation at the agency to reach high fidelity on this item in terms of personnel and the prevailing philosophy."
  - "I think the basic philosophy of the agency, the leadership team, the willingness to do the base work, to take the time because it is about improved outcomes for the clients. That is the bottom line here…. They're really committed to client focus. It just fits. Problem solving, solution focused, the best evidence makes sense. (HF2, 24M Program Leader Interview)"
  - "This observer finds an organization with extraordinary commitment and determination to serve dually diagnosed (DD) clients with the best services. Interviewees at multiple levels appear to embrace the integrated approach to treatment and basic principles of the model."
  - "The agency's adoption of a recovery philosophy is very much conducive to the implementation of the IDDT model." In addition, LF3 indicates: "The program philosophy (G1) articulated at baseline assessment by all parties included elements generally supportive of evidence-based practices, in that there is a strong commitment to improving clinical treatment services to the seriously mentally ill consumer."
- Concept of culture
  - "Representatives from supervised housing communicate regularly with team members as needed, and are receptive to learning and applying IDDT principles."
  - "Those addiction specialists are supervised by addiction managers, but they are then inserted into different teams and agencies throughout the whole agency, which involves a lot of different sites. So they are like centrally managed, but they are actually working day to day on other teams spread throughout the agency. (LF2, 6M Implementation Monitor Summary)"
  - "Clinicians report that when involved, professionals from other systems, e.g., child and family services, will attend a team or other inter-disciplinary meeting to collaborate on clients' treatment. In general, it was reported that there is good communication and easy access among the various disciplines. (LF1, 24M Fidelity Report)"
  - "openness of the sharing of ideas and working as a team"
  - "you never feel like you're hanging out there alone and I think the staff has really supported each other,"
  - "the strengths I think are the team itself, "

- "at the 12M they had an intact team that had basically all done training together and had really sort of been together fairly cohesive team."
- "So what's going well however is this team is dedicated to taking care of business and so they cover for each other."
- "Well, the program leaders I think…. we have tried to meet regularly to be very much involved, to define our roles, to encourage communication, top down bottom. I think there is a lot of teamwork. (HF2, 18M Program Leader Interview)'

Leadership

- Role of leadership
  - "The impetus for high fidelity at this site is the agency and program leadership," while "many of the strategies employed to increase fidelity were developed in the IDDT EBP leadership meetings."
  - "A good foundation at the agency was due to the leadership's knowledge and involvement."
  - "Let's do it in a way that helps the team understand why and what the reasons are behind it and so it encourages them and gives them the tools to do it."
  - "The practitioners have played a part in developing and carrying out strategies. Several practitioners serve on the leadership committee. They were chosen for this committee because they showed interest in the EBP. Any practitioner that expresses interest in the IDDT EBP seems encouraged to participate in planning process. (HF2, 24M Implementation Monitor Summary)"
  - ""The team leader/program leader and the Chief Operating Officer have developed an impressive document they issued to organize team meetings and structure consumer staffing,"
  - ""there continues to be plans to develop a new comprehensive assessment protocol/tool that will incorporate IDDT principles agency-wide."
  - "The Program Leader got it and was helping them get it before I ever showed up."
  - ""Without a knowledgeable, available, and somewhat charismatic team leader to guide team members through everyday problems and thereby
  - "At times, the leadership team members will negotiate alternatives to a recommendation if the proposal seems unrealistic. This negotiation is significant because the leadership team has excelled with following through on the initiatives they agree to implement. Lead by the executive director, they are able to weigh what they can realistically accomplish and what will have to wait, or be modified. (HF2, 12M Implementation Monitor Summary) Chief Operating Officer who is the champion of the model and herself a skilled clinician and supervisor. Her level of involvement, particularly in the planning and implementation stages was high, and she continues to be involved at this time in a more collaborative way with the CSP Coordinator. (HF3, 6M Implementation Monitor Summary)
- Stakeholders
  - "The leadership team members seem to take the trainer's recommendations, weigh them, and implement them when possible"
  - "The consultant/trainer was able to take what was essentially a pretty dry recitation of the model and really bring it to life in a way that connected the people very well."
  - "he's become very comfortable with the material and good at kind of knowing his audience and what do they need and how well these people learn best."
  - "Involving all different parties has been real beneficial,"
  - "In ongoing steering committee or leadership team meetings, this is where we really have laid the ground work for setting up the design and structure of the IDDT program."
  - "They are truly thinking about this steering committee as a way to involve others in the community through the implementation project."
  - "Due to the abundance of skill and energy at the program leader level and continued Board involvement and support, the program is humming along."

- "They've got a good county board commitment and good organizational board top leadership, the top CEO is very much supportive of it and I think their providers are pretty right for it as well."

Adaptability

- Consensus and buy-in
  - "strongly committed to build high fidelity integrated dual diagnosis programming within this agency."
  - "The willingness to make change; the openness to feedback that the agency has had, and the enthusiasm by staff and administration."
  - "A fidelity plan was developed following the 6M site visit and recommendations, and has driven activities to improve the program since that time."
  - HF3's response to improving fidelity was following "a blueprint" while "reaping the success." HF1 acknowledged at the baseline time point: "Buy-in at both board level and provider level and some of the managers,"
  - while at the 18M time point, "everyone's excited about the program and what we're doing."
  - "I don't think without the staff buy-in and without their work and energy and imagination and practical ideas and follow-through, that's absolutely essential."
  - "They need to buy in that this, will overall make their job easier and will make the outcomes of their clients; they have to believe that."
  - "Well, the first thing they did was, well, just say this was something we're going to do, which kind of doesn't leave people a lot of choice, you either buy into it or leave."
  - "The consumers have played certain key roles in the IDDT implementation," or "having people come in who tell me who are safe enough and secure enough to say this is what's going on….that's happening far more now than it used to." A 24M fidelity report at HF3 reflected "client participation in goal setting and feedback."
  - "The last steering committee that we had…. for the purpose of reviewing the six-month, fidelity was as good as any we've had yet with real strong consumer participation. It was the second time across the steering committees where I felt the consumers who were on the committee really stepped up and made a meaningful contribution to where we were going in terms of wanting to. (HF1, 6M Trainer Interview)"
- Changes in agency practices
  - "The most effective strategy that this agency has used so far has been to really structure the agency so that they could institutionalize change."
  - "So it was a kind of two-fold thing between the ongoing competence with skills from the direct service staff as well as the agency structure to support the practice."
  - "The culture of the way they do business is now built into those intensive teams: we stage, we do monthly staffing reports, we strategize by stage, we do family work, we do groups."
  - "A significant effort has gone into the overhaul of agency documentation to incorporate IDDT language and concepts."
  - "At the 12M point they had gone out and basically recruited staff specifically who wanted to do the model as opposed to assigning staff. They had changed the team leader role so that the program leader/team leader was going to be the sole focus for the IDDT initiative both from a management standpoint and from a team leader standpoint. They did a real nice job getting the right people in there because that group was able to talk the talk better than the training group before they'd been trained. (LF1, 12M Trainer Interview)"
- Changes in practitioner behaviors
  - "There is improvement noted in the consistency with which motivational stage is documented and interventions used, in the charts reviewed,"
  - "it's sort of a natural growth curve I guess for them in terms of just getting better at what they do by continuing to do it."
  - "Once case managers see consumers improve with the new programming, once they feel like they have some tools to be able to use with dual diagnosis consumers that are

effective, there's a contagious effect, and I think that's been real positive that it builds on its own momentum so that they are more motivated to use the skills that are taught in IDDT and that helps sustain the practice. (HF2, 24M Trainer Interview)"
    - "It's like they had a moment of clarity or something and harmonic convergence is upon them and they're starting to do some of that stuff."
    - "One of the things that's emerged is that the principles of what's going on here have sort of taken on more importance than the personalities involved,"
    - "They're now able to engage clients that used to be seen only in emergency/crisis situations, cases that they "thought were hopeless."
- Clear role delineation
    - "a harmonic convergence allowing people to bend their roles with an appropriate twist."
    - "She offers the structure and he sets the clinical direction," and "your machine, your operation here is cranking and I think it's because each of the three of you brings something to the table."

Knowledge

- Investment in training
    - "the managerial championing, the managerial investment because really that is the commitment to training,"
    - "every person on the team has gone to every training…..I credit a lot of their success to consistency because they're all hearing the same thing everywhere they go."
    - "we certainly have tried to pay attention with people coming on board, making sure that they have the orientation, and that we send them to training,"
    - "we developed a weekly study group of IDDT material in addition to formal training."
    - "The consultant/trainer estimates that most team members have had upwards of 50 hours of IDDT training at this point. Due to this reinforcement around the skill sets such as motivational interviewing and stage-wise treatment, the team is solid clinically with respect to the model. (HF3, 6M Implementation Monitor Summary)"
    - "Senior administrators report that agency-wide cross training, workshops, and in-services seminars carrying CEUs will be used to both refresh clinicians already trained and further disseminate IDDT in the agency."
    - "The Team Leader, dual diagnosis Program Manger and the Project Director have provided training for new hires (e.g., the new vocational specialist) and ongoing refresher and booster training both for the SAMI ACT team and to the wider agency staff. Training the new team member involved shadowing experienced staff, co-facilitating group, using the Implementation Resource Kit (IRK) manual, viewing the IRK videos, and meeting with the team leader and dual diagnosis program manager on an individual basis. (HF1, 12M Fidelity Report)"
- Mastering skills
    - "The team continues to develop, understand, and truly move toward mastery of the model's important skill sets"
    - "Clinicians appear to have a good grasp of motivational techniques and descriptions of activities with clients show solid understanding and appropriate use of the interventions."
    - "It was evident that many of the case managers understand the model and are able to integrate the model into their daily work."

Priority of supervision

- "In the team meetings…..the opportunities to reinforce the language, the structure and what have you are happening ….so I think that's just a huge piece of it and the clinicians rally by the model."
- "All team members participate in daily 1-hour team meetings led by the team leader, receive individual supervision from either the team leader or the dual disorder program manager for 1 hour per week, meet as a team once per month for 2 hours for rotating client. The team leader receives regular structured individual supervision from the dual disorder program manager.

Supervision is focused on clinicians' skills/professional development in IDDT service delivery (HF1, 12M Fidelity Report)"

Martinez et al. 2014

No quotes on facilitators were presented in this article

Murray et al. 2014

No quotes on facilitators were presented in this article

Owczarzak et al. 2011

Intervention level

- Fit
    - "[The SISTA intervention] fit very well. But most of the trainings [prevention interventions—JO] that I was providing at first were just single trainings, you know, the DEBI became a multisession [intervention] for one. It was more interactive in that, normally, when I'm doing an HIV 101 type of training, I am standing there just pushing information off on people. I may poll people to ask them their opinions about different things, but with DEBIs, [there are] all types of interactive exercises that people can—they don't get lost. [Female, direct service staff, 10 years with agency]"

Organizational level

- Broad support within the organization
    - "My supervisors have always been like, "How are things going? Are we implementing it? Have you had to make any adaptations? What is working and what is not working?" [They] do a pretty good job of it but it's because . . . they both have been doing HIV work in excess of 15 years, so they have been trained in the models as well. So they know they get [understand—JO] them and they follow up: "Is this working? Oh, OK." [Male, direct service staff, 4 years with agency]"
    - "You gotta send people [to training]—not only your outreach staff—but you got to send managers, or people that are going to supervise" (Female, direct service staff, 14 years with agency)."
- Flexibility
    - "I have to say that . . . the President and the Vice-President of this here organization . . . allow me to the full capacity to be able to serve them without . . . stepping in. They do guide the program. They will oversee the parts of what it takes in that part of management, but as far as delivering with fidelity and different things, they never hinder, and they really allow me to be able to go and do my thing . . . They're really good. [Female, direct service staff, 8 years with agency]
- Guidance
    - "I actually think that the State does a really good job with that. What has happened in the past, when we've had a training from the State, they will bring the trainers in . . .We go through the training. Then, a couple things can happen. [Sometimes] there is a specific group that is trained, and they're followed on the State level and the State will support us—when we got trained for VOICES/VOCES they made extra money available to buy a variety of condoms because a lot of people can't go out and buy a whole bunch . . . They've supported us in either having follow up meetings. They've had follow up conference calls, which can be helpful. [Our grant monitor has] CBO meetings, and every other month all of the folks who are funded by the State meet either in Madison or in the Milwaukee area. We get together and we talk about what we're implementing, how we're implementing it, what's really going well, what's not going well. Something that's not working for us may or may not be working for somebody else. And if [our grant monitor] keeps hearing there's something that really isn't working for everybody, she'll [say], "Ok, well we've got to do something about that," versus if somebody is doing it well and then there's one person just struggling, then that's another

opportunity to fit people together so that they have support from their other peers out doing HIV. [Female, Director, 1 year with agency]"

Program level

- Accessible facilitators
    - "We kind of checked in with the trainers and the trainers really kind of encouraged [us] to follow as closely as we could . . .We took that into mind as best we could adapt it to our population. We knew that we didn't want to exclude women because we thought that would not send the right message from our program. So we knew we wanted to include women, so that was something that we had to stray from. We looked at the initiative and said, "This is what we see as the core—these are the most important things—and we are going to do those things." Some of these other things that they suggested don't seem 100% vital to it and they don't fit for us anyway. So, we are just going to change those but we felt like we really kept true to the core of the initiative. [Male, Associate Director of Prevention, 1 year with agency]"

Owczarzak 2012

Provider perspectives on the need for EBP

- Seeing the value of EBP
    - "Once upon a time, prevention could not be measured. Now that we have these evidenced-based programs, we are seeing that prevention efforts can actually be measured and that is really important in many aspects, especially with program funding because a lot of people, in the past, or a lot of agents or government agencies didn't necessarily want to fund prevention programs because there was nothing concrete, no concrete information to say that they did in fact work. And so I think that having evidence-based prevention programs is a milestone toward proving that prevention works. [Female, 30s, director-level, 5 years with agency]"
- Fit with agency mission and identity
    - "[Thirteen, fourteen years ago], we didn't employ a DEBI. There was no DEBI. The woman who ran the program was called the 'Condom Queen', so I'm imagining that the gist of her intervention was condom demonstration. And fourteen years ago, that's what a lot of people did. But when more energy was put into [it], we need to have an evidencebased model . . . It's not enough for people to feel better. They have to really have gone through some change and know how to implement the change. [Female, 50s, director level, 20 years with agency]"
- Prevention efforts were worthwhile
    - "My motivation [for attending a DEBI training] really was to say, 'Wow! Here's a scientifically proven program that can be adopted and implemented in our area and I wanted to find out more about it . . . ' It was scientifically based. Here's probably one of very few programs that is actually scientifically proven to work to reduce the spread of the disease. Really, that is what we are here for and what we are doing . . . Going out and doing outreach: is that really the best use of our time? Or, here on the other hand, we have a program that's proven so why are we not already doing that program? [Male, 30s, direct service provision, 5 years with agency]"

What works in HIV prevention

- Seeing that the program appeared to be working
    - "Well, we do risk behaviors assessments if they come back a second time and compare those; that's like solid production of outcomes. I think that for me it's also when they bring other people to [the program]. It's like my peer educators that can go into the community and do presentations on safer sex. Being able to see them to do that: how many youth that they educate doing small group discussions, or larger group discussions for that matter, on safer sex, HIV and AIDS. For me that's a success. [Male, 30s, direct service provision, 8 years with agency]"

- "Whatever their life circumstance is, or why they're there and sometimes it's mostly referred or whatever, that they come in and they sit down [and in] the first few minutes [they're like], 'I really don't want to be here', or they give you excuses, [such as], 'I need to be looking for a job. I need to be doing this. I really shouldn't be here. What does this have to do with . . . being unemployed?' But those are the ones that show up after that group, in session number one, that they may show up for all five sessions. I find that very rewarding for someone that initially doesn't want to be there or has self-doubt why they're here, and then, after telling them, we're going to see, you're going to see why you're here – to learn skills. Because there's a lot of skills building within SISTA and we're talking about learning to be assertive versus being non-assertive or being aggressive. And that also goes along and skills for everything – jobs and anything. It doesn't just have to be within a relationship. They learn. [Female, 50s, direct service staff, 8 years with agency]"

Payan et al. 2017

Initiation

- Personnel capacity
  - Quality improvement
    - "[We] have previously looked at the [hypertension] rates by site, but not by provider. We hope to do that this year as part of our 2014 quality control goals."
- Technological capacity
  - Updated equipment to systematically capture data
    - "It's improved the quality of patient care and preventive medicine. And it's assisted us, the enhancement of our electronic medical record . . . in a way that will be enduring. It's not temporary."

Planning

- Professional developmental capacity
  - Activated medical directors who promote updated guidelines and practices
    - "Adult hypertension is a silent killer. . . . We currently identify acute and symptomatic, but we would like the clinical intervention to draw out the non-symptomatic and non-acute preventive cases and expand our existing electronic medical record capacity."

Pemberton et al. 2012

Implementation of MIP

- Agency leadership support
  - "Staff have [a] strong commitment to MIP clients. There is adequate supervision and support of MIP staff. Upper management is open to staff ideas to ensure program effectiveness....The Executive Director [is] attentive to staff needs and our suggestions to improve agency services."
  - "We get a lot of support from each other and from the Supervisor. We also have support from the Medical Director. We do both individual and group case conference two times a month. We discuss cases individually and as a group."
  - "Our Team has free range to do what is needed for the intervention. The agency allows funds to be spent as needed and assume that they will get it back to us. Finance is never a barrier in that area."
- Major successes with MIP Implementation
  - Positive client-level changes
    - "MIP helps client realize health is important. When clients are using drugs they forgot about themselves, so MIP helps raise awareness of needs. Eventually there are rewards. for example—when clients report that they have reduced their intake of drugs or that they are using condoms to protect themselves during intercourse, or when they say they are taking medication as mandated--things that they did not do before—to me that is an accomplishment."

        - "What has worked well with MIP implementation is that participants are empowered and see that they have the capacity to improve their lives. MIP participants have increased self confidence."
        - "In one case a client survived a domestic violence incident after being violently physically hurt. The participant did not relapse; she filed a police report and continued with MIP. [She] received medical and mental health services through MIP and is doing well."
    - MIP staff capacity, commitment and teamwork
        - "There was a lot that went into implementing the program-- putting a great team together, being able to fully understand the intervention, and being committed to the fidelity…. The team did so successfully."
        - "MIP staff work really well and [as] a unit. We are a close knit group. We have a great Coordinator that advocates for MIP staff."
        - "We do whatever it takes. The client gets the complete service. Staff work overtime if need be and staff hours are settled later. Many times clients are accompanied to medical and service appointments. This proves to the client that [agency] has a strong commitment to assist them and will not defraud or deceive them in any way.
    - Effective intervention components benefit participants
        - "I've had a lot of successes in MIP. I honestly believe it is the case management piece [in MIP] because you can get counseling anywhere…."
        - "Participants are "pleasantly surprised" with all of the MIP services that they are offered and seem to appear more secure and calmer with the process once the program is thoroughly explained to them."
        - "It is good to know that MIP is something that we can use and that it is really working… I think that they [clients] respond to what they have been given…at different times, at different levels. [MIP] meets the client at their pace."
        - "MIP is excellent. The participants accept it. MIP is tremendous."
    - Use of incentives
        - "The incentives have worked well in terms of recruitment and retention of clients."
        - "[Agency] offers significant incentives at the beginning of the program and even better at the end. [MIP clients] get gift cards of $ 20 for the 1st session (Induction session), $10 per subsequent sessions and $50 for the booster session. We give gift cards that have been suggested by participants and gift certificates to restaurants."
        - "Clients are offered incentives for a referral; persons are screened to see if eligible for the [MIP] program. We screen them when we test them for HIV."
        - "[Agency] sets aside weekly medical appointments for MIP participants (free of charge services). [They] receive on-site medical services--.family planning, screening and testing for HIV, comprehensive care clinic, mental health services, OB/GYN, youth services, pediatrics and dental care."
- MIP Program Implementation Support
    - Helpfulness of funder
        - "Project Officers are always there to provide assistance every time I call. [They] always have a response and a way to help me see things or suggest ways to help me solve the issue. [The] tools are out there and the people I have had the opportunity to work with have guided me and helped me."

Petrescu-Prahova et al. 2016

Implementation themes

- Organizational support
    - "When someone comes in and they're new… they [the staff] introduce them [the new person] to me… or they hand them a schedule. They'll show them around and suggest different classes." Female instructor, 2 years experience

    - "I started out in a small room and my class grew, and so they moved me into the gymnasium… if we need equipment, we can get it rather quickly." Female instructor, 4 years experience
    - "…there are always substitutes available if I'm unable to teach my one class a week. The YMCA [works hard to] acquire instructors that would be adequate for the classes." Female instructor, 5 years experience
- Instructor training
    - "…it was very well organized. The material was covered… very thoroughly. We had ample opportunity to ask questions… There was plenty of time to make it more personal for us, instead of just the blah, blah textbook." Female instructor, 1 year experience
    - "We sat down with a PowerPoint and went over the background of EnhanceFitness… I think it's very important to know… to be told the reasons why and the evidence behind the why was really cool." Female instructor, 1.5 years experience
    - "What I found helpful with the whole program is that it's structured, yet you have room for initiative. In the cardio section you can incorporate stuff from your background and from their background… when it comes to the weights, it's structured and so you can't go wrong." Female instructor, 4 years experience

Maintenance themes

- Organizational infrastructure for program delivery
    - "I think that having a consistent time and space for it, we find that that really helps programs to be successful." Female staff, 6 years experience
- Match between EF and Y Mission
    - "I'm a huge advocate of EnhanceFitness… Here at the Y we have three focus areas. We have youth development, healthy living, and social responsibility. Without a doubt, EnhanceFitness definitely hits two of those three…" Male staff, 1.5 years experience
    - "Becoming certified in EnhanceFitness allowed me to really connect with the members, especially our senior group. To me it's very important, especially during the day here that we have our seniors enjoying their time at the Y with us. EnhanceFitness has definitely helped us do that here." Male staff, 1.5 years experience
- Champions
    - "I'm a cheerleader for every program that I have here." Female staff, 2 years experience
    - "There were several people who were walking advertisements for the class, they would tell one another in class the things that they were now able to do that they
    - couldn't before." Female instructor, 1.5 years experience
    - "They didn't fill my position for quite some time, and so I think that not having a point person who was enthusiastic and passionate in making that a priority was a huge reason why it kind of fell through." Female staff, 3 years experience

Pinto et al. 2015

Costs

- Staff time and precollaboration
    - "No one thought that this (staff time) would be a big burden so we moved forward."
- Staff time and future collaborations
    - "the research is pretty easy going in terms of time commitment for staff"
- Financial costs
    - "this collaboration was funded by the study and the CBO just carried it out. This is helpful because the CBO doesn't have money in the budget to do this type of collaboration."

Benefits

- Benefits and precollaboration

    - "post-diagnosis can be very important piece in an individual's life because it provides an important teaching moment which can help change behavior to reduce risk for recurrence and co-morbidity."
    - "obvious benefits for patients who participated because the study promoted health and provided patient support."
    - "the CBO funds research and research is of interest to survivors, the type of data gained would be used to better serve survivors (this is a "win–win"), the RTR volunteers would also benefit and learn how they can better assist and help survivors."
- Benefits during collaboration
    - "this was a good opportunity to provide a program to survivors that they could participate in."
    - "RTR program benefitted because there was deeper involvement from the volunteers and that the collaboration helped to revitalize the RTR program."
    - "excites staff and spills over to the volunteers."

Match with CBO's mission and goals

- Match during collaboration
    - "shared vision was the most important: both groups had vested interest in survivorship. Both parties were totally on the same page here. Both wanted to know what is in the best interest of survivors and what do they need that they don't even know they need."
    - "had the same values as the study. Being physically active is important to the CBO. This was key to the partnership."
- Match and future collaboration
    - "must coincide with what they do and be in line with the goals and the mission of the CBO."
    - "mention the appropriate focus of the research study and that it is in line with our CBO's goals."
- Match with champions at the CBO
    - "without this, the CBO could easily take the collaboration off their plates."

Communication

- Communication and precollaboration
    - "communication was vital and although there was a glitch in the communications early on, this was resolved"
- Communication during collaboration
    - "The study staff was excellent in terms of communication. Communication was warm and inviting. There was always good communication, via phone, email or text."
    - "There was good, open communication to determine the needs of the study and ways to achieve goals. It was a pleasure to collaborate."

Ramanadhan et al. 2012

Program supports for implementation and adaptation

- Training
    - "One of the pieces that usually comes with an evidence-based program is that there is a training that goes along with it so you don't have to come up with a training for people. And also, an evaluation that comes up with it. Because, coming up with your own evaluation is a killer." —Participant 5 (Focus Group, Boston)
    - One of the grants that we've got.what they did was give you some intensive training on the background of how to do it . They had all the evidence for you, and gave you the tools to work with it so that you could modify and adapt it to your situation and take it. It helped.—Participant 10 (Focus Group, Worcester)

Collaborative technical assistance

- "One of the difficulties was, they have a very controlled population (but) . our population was very varied . So that was really difficult to quantify. You could see the result, but, it wasn't as easy

to quantify the way they did it in their report. So, it was a big drawback. But I had to call some of the guys . and they were very helpful to tell me what to do." —Participant 11 (Focus Group, Worcester)"

- We (went) back to the people who developed the curriculum . (to) say, "This doesn't work for our population, this age. We have to drop it down." And they said, "You know, as long as there was this much knowledge that they already had, then you could drop it down."—Participant 5 (Focus Group, Boston)
- When you demand all of this sort of technical stuff, we need more assistance from the funders, from the hospitals, from scientific-based community, whatever it is.so that we can go and do the work. We need your partnership more than just, you know, plopping money on us" .— Participant 5 (Focus Group, Boston)

Perceived benefits of using proven programs

- It has the accreditation, the American Diabetes Association stamp behind it, so I certainly feel that I'm comfortable and, um, you know, that it's got that background." —Participant 12 (Focus Group, Lawrence)
- I think it just lends credibility to what you're trying to do. With funders, whether it's your…your current funder, or if you're seeking new funding, . . there's a credibility behind it when you can say that you use evidencebased programs." —Participant 13 (Focus Group, Lawrence)

Thomas et al. 2014

No quotes on facilitators were presented in this article

Vanderpool et al., 2011

EBI is achievable

- 'seemed doable or interesting'

Veniegas et al. 2009

Maintenance and evolution phase

- Supporting organization change and institutionalization
    - "We are continuously looking to expand the services that we offer."
    - Fit "with all the care services that we have, case management, education, mental health, transportation, housing, food bank."

Appendix B5: Adapted checklist for quality appraisal of cross-sectional findings

Items		Yes	No	Not applicable	Unclear
Introduction					
1	Were the aims/objectives of the study clear?				
Methods					
2	Was the study design appropriate for the stated aim(s)?				
3	Was the sample size justified?				
4	Was the target/reference population clearly defined? (Is it clear who the research was about?)				
5	Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?				
6	Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?				
7	Were measures undertaken to address and categorise non-responders?				
8	Were the risk factor and outcome variables measured appropriate to the aims of the study? (e.g. could you see the questionnaire?)				
9	Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously?				
10	Is it clear what was used to determined statistical significance and/or precision estimates? (eg, p values, CIs)				
11	Were the methods (including statistical methods) sufficiently described to enable them to be repeated?				
Results					
12	Were the basic data adequately described?				
13	Does the response rate raise concerns about non-response bias?				
14	If appropriate, was information about non-responders described?				
15	Were the results internally consistent?				
16	Were the results for the analyses described in the methods, presented?				
Discussion					
17	Were the authors' discussions and conclusions justified by the results?				
18	Were the limitations of the study discussed?				

Other					
19	Were there any funding sources or conflicts of interest that may affect the authors' interpretation of the results?				
20	Was ethical approval or consent of participants attained?				

Quality (circle): **High** **Medium** **Low**

Reviewer: _____

Date: _____

Study: _____

Rationale and comment:
_____
_____
_____
_____
_____
_____

Appendix B6: Adapted CASP and JBI checklist

Checklist questions	Yes	No	Not applicable	Unclear
Was there a clear statement of the aims				
Is a qualitative methodology appropriate?				
Was the research design appropriate to address the research aims?				
Was the recruitment strategy/sampling appropriate to the aims of the research?				
Was the data collected in a way that addressed the research issue?				
Has the relationship between researcher and participant been adequately considered?				
Have ethical issues been taken into consideration?				
Was the data analysis sufficiently rigorous?				
Is there a clear statement of findings?				
Do the conclusions drawn in the research report flow form the analysis or interpretation of the data?				

Quality (circle): **High** **Medium** **Low**

Reviewer: _____

Date: _____

Study: _____

Rationale and comment:

_____

Appendix B7: Less significant recommendation themes

Main Themes	Times cited	Citation source
Recommendations for funders and collaborators		
Assess fit with other programming	1	(Belza et al., 2015)
Access cultural and demographic needs of the target population	1	(Belza et al., 2015)
Assess start-up and on-going costs and offsetting funding/revenue.	1	(Belza et al., 2015)
Explore adoption among majority and laggard adopters and compare to early adopters	1	(Belza et al., 2015)
Explore influence of adoption on implementation and maintenance	1	(Belza et al., 2015)
Explore policy approaches to revenue development	1	(Belza et al., 2015)
Include practitioners in the design of EBIs	1	(Flores et al., 2016)
Availability of other sources of funding and support in the external environment when implementation funding is ended.	1	(Hunter et al., 2017)
Invest in the relationship and the partnership	1	(Pemberton, 2012)
Enable and support adaptation of EBIs.	1	(Pemberton, 2012)
Emphasise benefits to TSO	1	(Pinto et al., 2015)
Continuous communication	1	(Pinto et al., 2015)
Identify champions	1	(Pinto et al., 2015)
Include TSO in needs assessment	1	(Pinto et al., 2015)
Recommendations for research and practice		
Service providers also suggested the need for clearer delineation of responsibilities related to the process of DEBI adaptation, especially given the time, staffing, and resource constraints of agencies	1	(Owczarzak & Dickson-Gomez, 2011)
To common understanding of the role of evidence in program planning and delivery.	1	(Ramanadhan et al., 2012)

Communicate modifications	1	(Lundgren et al., 2011)
Develop measurement tools of fidelity	1	(Lundgren et al., 2011)
Recommendation for practitioners		
Programs must consider community capacity	1	(Amodeo et al., 2011)
Willingness of organisations to engage in EB	1	(Lundgren et al., 2012)
Incorporate EBP in all levels of organisational infrastructure	1	(Murray et al., 2014)
Champions	1	(Petrescu-Prahova et al., 2016)
Invest in the relationship with funder	1	(Pemberton, 2012)
Create a culture of learning in the organisation	1	(Pemberton, 2012)
"Outreach to adolescent health care providers by program planners is recommended"	1	(Feutz & Andresen, 2013)
Understand target population	1	(Feutz & Andresen, 2013)
Introduce incentive for population to participate	1	(Feutz & Andresen, 2013)

Appendix B8: Quality Appraisals of qualitative studies

Study	Was there a clear statement of the aims ?	Is a qualitative methodology appropriate?	Was the research design appropriate to address the	Was the recruitment strategy/sampling appropriate to the aims of the research?	Was the data collected in a way that addressed the research issue?	Has the relationship between researcher and participant been adequately considered? research issue?	Have ethical issues been taken into consideration?	Was the data analysis sufficiently rigorous?	Is there a clear statement of findings?	Do the conclusions drawn in the research report flow form the analysis or interpretation of the data?	Quality assessment
Allicock et al. 2012	X	X	X	X	X		X	X	X	X	Medium
Amodeo et al. 2011	X	X	X	X	X		X	X	X	X	High
Belza et al. 2014	X	X	X	X	X		X	X	X	X	High
Collins et al. 2010	X	X	X	X	X				X	X	Medium
D'Ippolito et al. 2015	X	X	X	X	X		X	X	X	X	Medium
Demby et al. 2014	X	X	X	X	X				X	X	Medium
Dolcini et al. 2010	X	X	X	X	X		X	X	X	X	High
Feutz et al. 2013	X		X	X	X		X		X		Low
Flores et al. 2016	X	X	X	X	X		X		X	X	Medium
Gandelman et al. 2012	X	X	X	X	X		X	X	X	X	Medium
Honeycutt et al. 2012	X	X	X	X	X		X	X	X	X	High

Kegeles et al. 2015	X	X	X	X	X		X	X	X	X	High
Kimber et al. 2012	X	X	X	X	X		X	X	X	X	Medium
Lattimore et al. 2010	X	X	X	X	X			X	X	X	Medium
Lundgren et al. 2011	X	X	X	X	X		X	X	X	X	High
Maharaj 2010	X	X	X	X	X		X	X	X	X	Medium
Murray et al. 2014	X	X	X	X	X				X	X	Low
Owczarzak et al. 2011	X	X	X	X	X		X	X	X	X	Medium
Owczarzak 2012	X	X	X	X	X		X	X	X	X	Medium
Payan et al. 2017	X	X	X	X	X		X	X	X	X	High
Pemberton et al. 2012	X	X	X	X	X		X	X	X	X	Medium
Petrescu-Prahova et al. 2016	X	X	X	X	X		X	X	X	X	High
Pinto et al. 2015	X	X	X	X	X		X	X	X	X	Medium
Ramanadhan et al. 2012	X	X	X	X	X	X		X	X	X	High
Vanderpool et al. 2011	X	X	X	X	X		X		X	X	Medium
Veniegas et al. 2009	X	X	X	X	X		X	X	X	X	Medium

Appendix B9: Quality Appraisals of Quantitative Studies

Quality items	House et al. 2017	Hunter et al. 2017	Lundgren et al. 2012	Martinez et al. 2014	Thomas et al. 2014
The aims/objectives of the study were clear	X	X	X	X	X
The study design was appropriate for the stated aim(s)	X	X	X	X	X
The sample size was justified	X	X		X	X
The target/reference population was clearly defined (It clear who the research was about?)	X	X	X	X	X
The sample frame was taken from an appropriate population base so that it closely represented the target/reference population under investigation	X	X	X	X	X
The selection process was likely to select subjects/participants that were representative of the target/reference population under investigation	X	X	X	X	X
Measures were undertaken to address and categorize non-responders					X
The risk factor and outcome variables were measured appropriately to the aims of the study (e.g. the questionnaire could be seen)		X	X	X	X
The risk factor and outcome variables were measured correctly using instruments/measurements that had been trialled, piloted or published previously?		X	X	X	
It is clear what was used to determined statistical significance and/or precision estimates?		X	X	X	X
The methods (including statistical methods) were sufficiently described to enable them to be repeated		X	X		X
The basic data were adequately described	X	X	X	X	X
The response rate does not raise concerns about non-response bias		X	X		
If appropriate, information about non-responders was described					
The results were internally consistent	X	X	X	X	X

The results for the analyses described in the methods were presented	X	X	X	X	X
The authors' discussions and conclusions are justified by the results	X	X	X	X	X
The limitations of the study were discussed	X	X	X	X	X
Funding sources or conflicts of interest did not likely affect the authors' interpretation of the results	X	X	X	X	X
Ethical approval or consent of participants attained		X	X	X	X
Final quality assessment	**Low**	**Med**	**Med**	**Med**	**Med**

Appendix C (Chapter 5)

Appendix C1: descriptive information and predicted probabilities

Table A1: Description of all included care services regulated by the Care Inspectorate (as defined by the CI)

Service type	Definition
Care Home Service:	A service which provides accommodation, together with nursing, personal care or personal support, for persons by reason of their vulnerability or need this may include for: alcohol & drug misuse, blood borne virus, children & young people, learning disabilities, mental health problems, older people, physical & sensory impairment or respite care & short breaks.
Child Care Agency:	A service which consists of or includes supplying, or introducing to persons who use the service, child carers up to the age of 16, wholly of mainly in the home of the child's parents.
Day Care of Children:	A service which consists of any form of care on a non-domestic premises, supervised by a responsible person during the day for at least two hours a day on at least six days per year. The service definition includes nursery classes, creches, after school clubs and playgroups but does not include services which are provided as part of a school activity. Nor does it include activities where care is not provided such as sports clubs or uniformed activities such as Scouts or Guides.
Housing Support Service:	A service which provides support, assistance, advice or counselling to a person who has particular needs, with a view to enabling that person to occupy residential accommodation as a sole or main residence.
School Care Accommodation Service:	A service which is provided to a pupil by an education authority or the managers of an independent or grant-aided school for the purpose of the pupil being in attendance at a public, independent or grant-aided school; and which consists of the provision, in a place in or outwith the school, of residential accommodation. This may be take the form of mainstream residential accommodation, residential special accommodation or school hostel accommodation.
Secure Accommodation Service:	A service which provides accommodation for the purpose of restricting the liberty of children in residential premises where care services are provided as with regulations made under section 29(9)(a) of the Regulation of Care (Scotland) Act 2001. This may be to provide protection to the child who may have history of running away and whose health or wellbeing may subsequently be under threat.
Support Service:	A personal care or personal support service provided by arrangement made by a local authority or health body to a vulnerable or person in need. This does not include a care home service, a service providing overnight accommodation, adoption service, or fostering service.

Table A2: Distribution of included care services over sector

	For-profit sector	Public sector	Third sector	Total
Care Home Service	805	256	342	1,403
Child Care Agency	14	1	10	25
Day Care of Children	1,138	1,723	821	3,682
Housing Support service	227	180	654	1,061
School Care Accommodation services	20	7	36	63
Secure Accommodation service	1	1	3	5
Support Service	433	335	683	1,451

Table A3: Predicted probabilities over 3 levels of quality after generalised ordered logistic regression.

Quality of care and support (n=7,011)	**Low quality**	**Adequate quality**	**High quality**
For-profit sector	0.0153	0.4011	0.5836
95% confidence interval	0.0093 to 0.0213	0.3747 to 0.4275	0.5573 to 0.6099
Public sector	0.0045	0.3034	0.692
95% confidence interval	0.001 to 0.0081	0.28 to 0.3268	0.6686 to 0.7155
Third sector	0.0102	0.322	0.6678
95% confidence interval	0.0047 to 0.0156	0.298 to 0.3461	0.6438 to 0.6918
Quality of staff (n=7,007)			
For-profit sector	0.0169	0.4569	0.5262
95% confidence interval	0.0107 to 0.0231	0.4296 to 0.4842	0.4989 to 0.5535
Public sector	0.0038	0.3319	0.6643
95% confidence interval	0.0005 to 0.0071	0.3074 to 0.3563	0.6398 to 0.6888
Third sector	0.0121	0.3826	0.6053
95% confidence interval	0.0062 to 0.018	0.3576 to 0.4076	0.5803 to 0.6303
Quality of management (n=7,004)			
For-profit sector	0.0259	0.5144	0.4597
95% confidence interval	0.0182 to 0.0336	0.4867 to 0.5421	0.4321 to 0.4873
Public sector	0.0066	0.4212	0.5722
95% confidence interval	0.0023 to 0.0108	0.3958 to 0.4467	0.5468 to 0.5976
Third sector	0.0185	0.4443	0.5372
95% confidence interval	0.0114 to 0.0256	0.4186 to 0.4699	0.5116 to 0.5627
Quality of environment (n=5,099)			
For-profit sector	0.0087	0.476	0.5153
95% confidence interval	0.0041 to 0.0134	0.4437 to 0.5083	0.4829 to 0.5476
Public sector	0.0053	0.3698	0.6248
95% confidence interval	0.0004 to 0.0102	0.3432 to 0.3964	0.5983 to 0.6514
Third sector	0.0093	0.4425	0.5482
95% confidence interval	0.0022 to 0.0165	0.4085 to 0.4765	0.5143 to 0.582
Risk of services (n=6,891)	**Low risk**	**Medium risk**	**High risk**
For-profit sector	0.7991	0.1319	0.069
95% confidence interval	0.7793 to 0.8188	0.1146 to 0.1493	0.0569 to 0.081
Public sector	0.911	0.069	0.02
95% confidence interval	0.8955 to 0.9265	0.0549 to 0.0831	0.012 to 0.028
Third sector	0.8703	0.0847	0.0451
95% confidence interval	0.8543 to 0.8862	0.0708 to 0.0985	0.0345 to 0.0556

Table A4: Predicted probabilities of being issued complaints/inspection-requirements.

Complaints	Predicted probability	95 % CI	ORs	95 % CI
For-profit	0.2345	0.2128 to 0.2562	Reference group	
Public sector	0.0639	0.0513 to 0.0765	.1827634***	.1457376 to .2291957
Third sector	0.095	0.0806 to 0.1094	.2957996***	.2432594 to .3596876
Requirements				
For-profit	0.2667	0.2431 to 0.2902	Reference group	
Public sector	0.1618	0.1413 to 0.1823	.4935881***	.4117321 to .5917178
Third sector	0.2105	0.1886 to 0.2323	.7069104***	.5971361 to .8368649

Table A5: Cross tabulation of the four biggest service types across all quality and risk domains.

	Care home services			
Quality of care (n=1,327)	For-profit sector	Public sector	Third sector	Total
Poor quality	23 3.1%	4 1.6%	4 1.19%	31 2.34%
Adequate quality	371 50.07%	95 38%	99 29.46%	565 42.58%
High quality	347 46.83%	151 60.4%	233 69.35%	731 55.09%
Quality of staff (n=1,326)				
Poor quality	20 2.7%	3 1.2%	6 1.79%	29 2.19%
Adequate quality	397 53.65%	98 39.2%	104 30.95%	599 45.17%
High quality	323 43.65%	149 59.6%	226 67.26%	698 52.64%
Quality of management (n=1,327)				
Poor quality	30 4.05%	4 1.6%	6 1.79%	40 3.01%
Adequate quality	397 53.58%	113 45.2%	129 38.39%	639 48.15
High quality	314 42.38%	133 53.2%	201 59.82%	648 48.83
Quality of environment (n=1,326)				
Poor quality	17 2.3%	1 0.4%	4 1.19%	22 1.66%
Adequate quality	424 57.3%	103 41.2%	102 30.36%	629 47.44%
High quality	299 40.41%	146 58.4%	230 68.45%	675 50.9%
Risk of services (n=1,391)				
Low risk	490 61.71%	207 81.18%	280 81.87%	977 70.24%
Medium risk	157 19.77%	35 13.73%	34 9.94%	226 16.25%
High risk	147 18.51%	13 5.1%	28 8.19%	188 13.52%
	Day Care of Children			
Quality of care (n=3,568)	For-profit sector	Public sector	Third sector	Total
Poor quality	13 1.23%	4 0.24%	10 1.24%	27 0.76%
Adequate quality	494 46.6%	445 26.15%	311 38.59%	1,250 35.03%
High quality	553 52.17%	1,253 73.62%	485 60.17%	2,291 64.21%
Quality of staff (n=3,567)				
Poor quality	20 1.89%	3 0.18%	11 1.36%	34 0.95%
Adequate quality	529 49.91%	463 27.22%	381 47.27%	1,373 38.49%

High quality	511 48.21%	1,235 72.6%	414 51.36%	2,160 60.56%
Quality of management (n=3,567)				
Poor quality	32 3.02%	5 0.29%	22 2.73%	59 1.65%
Adequate quality	590 55.66%	662 38.92%	445 55.21%	1,697 47.57%
High quality	438 41.32%	1,034 60.79%	339 42.06%	1,811 50.77%
Quality of environment (n=3,568)				
Poor quality	8 0.75%	6 0.35%	6 0.74%	20 0.56%
Adequate quality	511 48.21%	588 34.55%	385 47.77%	1,484 41.59%
High quality	541 51.04%	1,108 65.1%	415 51.49%	2,064 57.85%
Risk of services (n=3,678)				
Low risk	981 86.51%	1,683 97.68%	736 89.65%	3,400 92.44%
Medium risk	106 9.35%	32 1.86%	60 7.31%	198 5.38%
High risk	47 4.14%	8 0.46%	25 3.05%	80 2.18%
		Housing Support Services		
Quality of care (n=996)	For-profit sector	Public sector	Third sector	Total
Poor quality	6 3.02%	1 0.58%	2 0.32%	9 0.9%
Adequate quality	73 36.68%	57 33.33%	136 21.73%	266 26.71%
High quality	120 60.3%	113 66.08%	488 77.96%	721 72.39%
Quality of staff (n=995)				
Poor quality	5 2.51%	1 0.58%	4 0.64%	10 1.01%
Adequate quality	89 44.72%	68 39.77%	172 27.52%	329 33.07%
High quality	105 52.76%	102 59.65%	449 71.84%	656 65.93%
Quality of management (n=993)				
Poor quality	7 3.52%	1 0.58%	3 0.48%	11 1.11%
Adequate quality	106 53.27%	81 47.37%	206 33.07%	393 39.58%
High quality	86 43.22%	89 52.05%	414 66.45%	589 59.32%
Quality of environment (no observations)				

Risk of services (n=1,057)				
Low risk	165 73.33%	143 79.44%	565 86.66%	873 82.59%
Medium risk	38 16.89%	31 17.22%	64 9.82%	133 12.58%
High risk	22 9.78%	6 3.33%	23 3.53%	51 4.82%
		Support services		
Quality of care (n=1,334)	For-profit sector	Public sector	Third sector	Total
Poor quality	8 2.12%	1 0.31%	3 0.47%	12 0.9%
Adequate quality	160 42.44%	104 32.1%	149 23.54%	413 30.96%
High quality	209 55.44%	219 67.59%	481 75.99%	909 68.14%
Quality of staff (n=1,333)				
Poor quality	9 2.39%	1 0.31%	4 0.63%	14 1.05%
Adequate quality	201 53.32%	122 37.65%	199 31.49%	522 39.16%
High quality	167 44.3%	201 62.04%	429 67.88%	797 59.79%
Quality of management (n= 1,331)				
Poor quality	12 3.19%	4 1.23%	6 0.95%	22 1.65%
Adequate quality	214 56.91%	147 45.37%	226 35.82%	587 44.1%
High quality	150 39.89%	173 53.4%	399 63.23%	722 54.24%
Quality of environment (n=437)				
Poor quality	0 0	0 0	1 0.59%	1 0.23%
Adequate quality	33 58.93%	66 31.13%	49 28.99%	148 33.87
High quality	23 41.07%	146 68.87%	119 70.41%	288 65.9%
Risk of services (n=1,448)				
Low risk	318 73.61%	287 85.67%	591 86.78%	1,196 82.6%
Medium risk	74 17.13%	39 11.64%	66 9.69%	179 12.36%
High risk	40 9.26%	9 2.69%	24 3.52%	73 5.04%

Appendix C2: Regression coefficients of all models

Quality of care

```
Generalized Ordered Logit Estimates                   Number of obs   =       7011
                                                      LR chi2(32)     =     780.71
                                                      Prob > chi2     =     0.0000
Log likelihood = -4488.0679                           Pseudo R2       =     0.0800

------------------------------------------------------------------------------
           care_qual | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------+--------------------------------------------------------
1                    |
               ordsec |
                   1  |   3.617589   1.428839     3.26   0.001     1.668092    7.845458
                   2  |   1.551432   .4831747     1.41   0.158     .8426293    2.856467
                      |
                  old |   .8780931   .6184372    -0.18   0.854     .2208223     3.49171
                  age |   1.028746   .0304255     0.96   0.338     .9708085    1.090141
           active_dum |    2.48884   1.924425     1.18   0.238     .5468014    11.32829
           child_care |   2.623745   2.894489     0.87   0.382     .3019149    22.80125
          NumberStaff |    .995739   .0018954    -2.24   0.025      .992031    .9994608
             children |   .6277419   .4902933    -0.60   0.551     .1358177    2.901389
housingsupport_dummy |   1.880857   2.604932     0.46   0.648     .1245852    28.39521
supportservice_dummy |   1.504408   2.062675     0.30   0.766     .1023986    22.10228
      carehome_dummy |   1.068072   1.268826     0.06   0.956     .1040897    10.95956
          school_dum |   851636.5   4.49e+08     0.03   0.979            0           .
            dementia |   1.114056   .5725881     0.21   0.834     .4068344    3.050679
                      |
        grade_change |
                   1  |   .0489929   .0127279   -11.61   0.000     .0294441    .0815205
                   2  |    2.36307   2.406158     0.84   0.398     .3211843    17.38597
```

3	.2890156	.1719283	-2.09	0.037	.0900663	.9274282
_cons	25.09901	39.60822	2.04	0.041	1.138657	553.2483
2						
ordsec						
1	1.664025	.12034	7.04	0.000	1.444116	1.917421
2	1.474978	.1036209	5.53	0.000	1.285246	1.692718
old	.4318115	.0844141	-4.30	0.000	.2943705	.6334234
age	1.044373	.0066945	6.77	0.000	1.031334	1.057577
active_dum	1.493337	.3504736	1.71	0.088	.9427313	2.365525
child_care	1.051557	.4380482	0.12	0.904	.4647795	2.379133
NumberStaff	.9930645	.001156	-5.98	0.000	.9908014	.9953327
children	.8973683	.2032698	-0.48	0.633	.57565	1.398888
housingsupport_dummy	1.969249	.9435612	1.41	0.157	.7699263	5.036768
supportservice_dummy	1.364967	.6511649	0.65	0.514	.5358607	3.476902
carehome_dummy	1.744381	.7734211	1.25	0.210	.731538	4.159545
school_dum	5.737416	3.398512	2.95	0.003	1.796859	18.31972
dementia	.8022336	.1385563	-1.28	0.202	.5718567	1.12542
grade_change						
1	.1674335	.0228994	-13.07	0.000	.1280637	.2189066
2	.4715262	.0558757	-6.34	0.000	.3737996	.5948027
3	.4293874	.0878028	-4.13	0.000	.2876023	.6410711
_cons	.6165617	.3297748	-0.90	0.366	.2161235	1.75894

Quality of staff

Generalized Ordered Logit Estimates — Number of obs = 7007; LR chi2(32) = 624.53; Prob > chi2 = 0.0000; Log likelihood = -4770.4883; Pseudo R2 = 0.0614

qual_staff	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1						
ordsec						
1	1.56036	.4134778	3.77	0.000	.7499584	2.370762
2	.3584772	.2820603	1.27	0.204	-.1943509	.9113053
old	.5366185	.6639625	0.81	0.419	-.7647241	1.837961
age	.0481668	.0275312	1.75	0.080	-.0057935	.102127
active_dum	.7248618	.7649833	0.95	0.343	-.774478	2.224201
child_care	-13.48848	979.7484	-0.01	0.989	-1933.76	1906.783
NumberStaff	-.0020215	.0029408	-0.69	0.492	-.0077853	.0037423
children	-.160296	.7128734	-0.22	0.822	-1.557502	1.23691
housingsupport_dummy	-13.45075	979.7487	-0.01	0.989	-1933.723	1906.821
supportservice_dummy	-13.52538	979.7487	-0.01	0.989	-1933.797	1906.747
carehome_dummy	-14.13483	979.7485	-0.01	0.988	-1934.407	1906.137
school_dum	.3588927	1414.548	0.00	1.000	-2772.104	2772.822
dementia	-.2628266	.565618	-0.46	0.642	-1.371418	.8457644
grade_change						
1	-2.916905	.2552869	-11.43	0.000	-3.417259	-2.416552
2	.979674	1.015763	0.96	0.335	-1.011185	2.970533
3	-1.636685	.4557055	-3.59	0.000	-2.529851	-.7435185
_cons	17.01345	979.749	0.02	0.986	-1903.259	1937.286

2						
ordsec						
1	.603499	.0700092	8.62	0.000	.4662834	.7407145
2	.3371037	.0674661	5.00	0.000	.2048725	.4693349
old	-.5826408	.1830049	-3.18	0.001	-.9413237	-.2239579
age	.0536388	.0061904	8.66	0.000	.0415058	.0657718
active_dum	.2471304	.2330696	1.06	0.289	-.2096777	.7039385
child_care	-.3758204	.4148514	-0.91	0.365	-1.188914	.4372735
NumberStaff	-.0030292	.0010036	-3.02	0.003	-.0049962	-.0010622
children	.0370988	.2109649	0.18	0.860	-.3763847	.4505824
housingsupport_dummy	.1782639	.4693637	0.38	0.704	-.7416721	1.0982
supportservice_dummy	-.1462852	.467938	-0.31	0.755	-1.063427	.7708564
carehome_dummy	-.0176229	.4392484	-0.04	0.968	-.878534	.8432882
school_dum	1.08317	.5621073	1.93	0.054	-.0185406	2.18488
dementia	-.0575563	.1702268	-0.34	0.735	-.3911947	.2760821
grade_change						
1	-.9503409	.1239546	-7.67	0.000	-1.193287	-.7073945
2	-.5843002	.1182296	-4.94	0.000	-.8160259	-.3525746
3	-.71315	.2131844	-3.35	0.001	-1.130984	-.2953163
_cons	-.437402	.5261457	-0.83	0.406	-1.468629	.5938247

Quality of management

Generalized Ordered Logit Estimates

Number of obs	7004
LR chi2(32)	541.98
Prob > chi2	0.0000
Log likelihood = -5104.49	Pseudo R2 = 0.0504

qual_man	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1						
ordsec						
1	1.470659	.3196049	4.60	0.000	.844245	2.097073
2	.3702209	.2313883	1.60	0.110	-.0832919	.8237337
old	.0011429	.6155718	0.00	0.999	-1.205356	1.207641
age	.0451738	.0224192	2.01	0.044	.0012329	.0891148
active_dum	.2447983	.7620621	0.32	0.748	-1.248816	1.738413
child_care	-.046742	1.074729	-0.04	0.965	-2.153172	2.059688
NumberStaff	-.0039073	.0016931	-2.31	0.021	-.0072257	-.0005888
children	.3135191	.7891894	0.40	0.691	-1.233264	1.860302
housingsupport_dummy	1.124893	1.36565	0.82	0.410	-1.551732	3.801519
supportservice_dummy	.5467217	1.345445	0.41	0.684	-2.090301	3.183745
carehome_dummy	.4720122	1.2185	0.39	0.698	-1.916204	2.860228
school_dum	13.83781	756.2878	0.02	0.985	-1468.459	1496.135
dementia	-.6698529	.4359013	-1.54	0.124	-1.524204	.184498
grade_change						
1	-2.853141	.2080409	-13.71	0.000	-3.260893	-2.445388
2	1.541346	1.010538	1.53	0.127	-.4392712	3.521963
3	-1.23886	.4276384	-2.90	0.004	-2.077015	-.4007037
_cons	2.950242	1.550329	1.90	0.057	-.0883464	5.98883

2						
ordsec						
1	.4680099	.0685499	6.83	0.000	.3336545	.6023652
2	.3212412	.0668904	4.80	0.000	.1901384	.4523439
old	-.4003983	.1786427	-2.24	0.025	-.7505316	-.0502649
age	.039026	.006131	6.37	0.000	.0270095	.0510426
active_dum	.0486828	.2304928	0.21	0.833	-.4030748	.5004405
child_care	-.2583895	.4025476	-0.64	0.521	-1.047368	.5305893
NumberStaff	-.0022101	.0009563	-2.31	0.021	-.0040843	-.0003358
children	-.0134476	.2039318	-0.07	0.947	-.4131467	.3862514
housingsupport_dummy	.2675281	.454912	0.59	0.556	-.6240831	1.159139
supportservice_dummy	.0252316	.4536396	0.06	0.956	-.8638856	.9143488
carehome_dummy	.1925452	.4258836	0.45	0.651	-.6421713	1.027262
school_dum	1.386507	.5425234	2.56	0.011	.3231803	2.449833
dementia	-.2726418	.1709459	-1.59	0.111	-.6076896	.062406
grade_change						
1	-1.015192	.1294059	-7.85	0.000	-1.268823	-.7615612
2	-.4255027	.1177498	-3.61	0.000	-.6562881	-.1947173
3	-.883682	.2270271	-3.89	0.000	-1.328647	-.4387171
_cons	-.4740285	.5122829	-0.93	0.355	-1.478085	.5300276

Quality of environment

Generalized Ordered Logit Estimates

Number of obs	5099
LR chi2(32)	349.14
Prob > chi2	0.0000
Log likelihood = -3510.4682	Pseudo R2 = 0.0474

qual_environ	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1						
ordsec						
1	.5119585	.4793881	1.07	0.286	-.427625	1.451542
2	-.0497575	.4368961	-0.11	0.909	-.9060582	.8065432
old	.3204709	.7762581	0.41	0.680	-1.200967	1.841909
age	.0915891	.0367722	2.49	0.013	.0195169	.1636614
active_dum	1.517168	.7982178	1.90	0.057	-.0473103	3.081646
child_care	-14.90339	3694.648	-0.00	0.997	-7256.279	7226.473
NumberStaff	-.0099754	.0104203	-0.96	0.338	-.0303988	.0104479
children	1.376886	1.201762	1.15	0.252	-.9785241	3.732296
housingsupport_dummy	1.523753	9488.146	0.00	1.000	-18594.9	18597.95
supportservice_dummy	-12.95929	3694.648	-0.00	0.997	-7254.336	7228.418
carehome_dummy	-14.4477	3694.648	-0.00	0.997	-7255.824	7226.929
school_dum	-.0511016	4186.933	-0.00	1.000	-8206.289	8206.186
dementia	-.2397758	.5573478	-0.43	0.667	-1.332157	.8526057
grade_change						
1	-2.743418	.3475525	-7.89	0.000	-3.424608	-2.062228
2	13.85418	638.2597	0.02	0.983	-1237.112	1264.82
3	-.830322	.8086726	-1.03	0.305	-2.415291	.7546473
_cons	16.47663	3694.648	0.00	0.996	-7224.9	7257.854

2						
ordsec						
1	.4650903	.0790323	5.88	0.000	.3101898	.6199908
2	.137697	.083688	1.65	0.100	-.0263284	.3017224
old	-.8593083	.1876587	-4.58	0.000	-1.227113	-.491504
age	.0232359	.0074236	3.13	0.002	.0086859	.0377859
active_dum	.0814776	.2641571	0.31	0.758	-.4362607	.5992159
child_care	1.229794	1.166721	1.05	0.292	-1.056938	3.516526
NumberStaff	.0032956	.0023765	1.39	0.166	-.0013622	.0079534
children	-.1186919	.2098222	-0.57	0.572	-.529936	.2925521
housingsupport_dummy	1.924667	1.658254	1.16	0.246	-1.325452	5.174786
supportservice_dummy	1.382002	1.189131	1.16	0.245	-.9486524	3.712656
carehome_dummy	1.551178	1.173367	1.32	0.186	-.74858	3.850936
school_dum	1.690782	1.19893	1.41	0.158	-.659078	4.040643
dementia	-.5361241	.1735847	-3.09	0.002	-.8763439	-.1959043
grade_change						
1	-.6237244	.1458968	-4.28	0.000	-.9096768	-.337772
2	-.7660555	.1443506	-5.31	0.000	-1.048977	-.4831336
3	-1.036217	.2742086	-3.78	0.000	-1.573656	-.4987783
_cons	-1.372131	1.221586	-1.12	0.261	-3.766395	1.022134

Risk of services

Generalized Ordered Logit Estimates

Number of obs	7335
LR chi2(32)	911.79
Prob > chi2	0.0000
Pseudo R2	0.1229

Log likelihood = -3254.8951

risk	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1						
ordsec						
1	-1.049211	.1077008	-9.74	0.000	-1.260301	-.8381213
2	-.583443	.090511	-6.45	0.000	-.7608414	-.4060447
old	.4334521	.2110812	2.05	0.040	.0197407	.8471636
age	-.0180398	.0085729	-2.10	0.035	-.0348424	-.0012372
active_dum	-.1379858	.3148894	-0.44	0.661	-.7551577	.4791861
child_care	.0672593	.6344598	0.11	0.916	-1.176259	1.310778
NumberStaff	.0080025	.0011696	6.84	0.000	.0057101	.010295
children	-.1438598	.2614877	-0.55	0.582	-.6563663	.3686468
housingsupport_dummy	.570841	.6888947	0.83	0.407	-.7793678	1.92105
supportservice_dummy	.7807085	.6873573	1.14	0.256	-.566487	2.127904
carehome_dummy	.8119581	.6585732	1.23	0.218	-.4788216	2.102738
school_dum	.6128083	.7412761	0.83	0.408	-.8400661	2.065683
dementia	.2299525	.1744474	1.32	0.187	-.1119582	.5718632
grade_change						
1	1.706355	.1309808	13.03	0.000	1.449637	1.963073
2	.9717695	.1359324	7.15	0.000	.705347	1.238192
3	-.2449884	.1618032	-1.51	0.130	-.5621169	.0721401
_cons	-1.758925	.7606758	-2.31	0.021	-3.249822	-.2680273

2						
ordsec						
1	-1.381424	.2028329	-6.81	0.000	-1.778969	-.9838786
2	-.490703	.1479122	-3.32	0.001	-.7806057	-.2008003
old	.4614046	.3090323	1.49	0.135	-.1442876	1.067097
age	-.03324	.0137416	-2.42	0.016	-.060173	-.0063069
active_dum	-.5714107	.3997949	-1.43	0.153	-1.354994	.2121729
child_care	-.2703299	1.034473	-0.26	0.794	-2.29786	1.7572
NumberStaff	.0045122	.0013453	3.35	0.001	.0018755	.0071489
children	.1349647	.3912695	0.34	0.730	-.6319094	.9018388
housingsupport_dummy	.4661248	1.109885	0.42	0.675	-1.70921	2.641459
supportservice_dummy	.6176765	1.108475	0.56	0.577	-1.554895	2.790248
carehome_dummy	.953676	1.06416	0.90	0.370	-1.13204	3.039392
school_dum	.962504	1.138753	0.85	0.398	-1.269412	3.19442
dementia	.386295	.2186784	1.77	0.077	-.0423069	.8148968
grade_change						
1	1.873533	.1616649	11.59	0.000	1.556675	2.19039
2	.882429	.1991498	4.43	0.000	.4921026	1.272755
3	-.3510296	.2654135	-1.32	0.186	-.8712306	.1691714
_cons	-2.455577	1.192812	-2.06	0.040	-4.793446	-.1177072

Complaints

```
Logistic regression                                     Number of obs   =       7358
                                                        LR chi2(16)     =    1216.61
                                                        Prob > chi2     =     0.0000
Log likelihood = -2314.3752                             Pseudo R2       =     0.2081

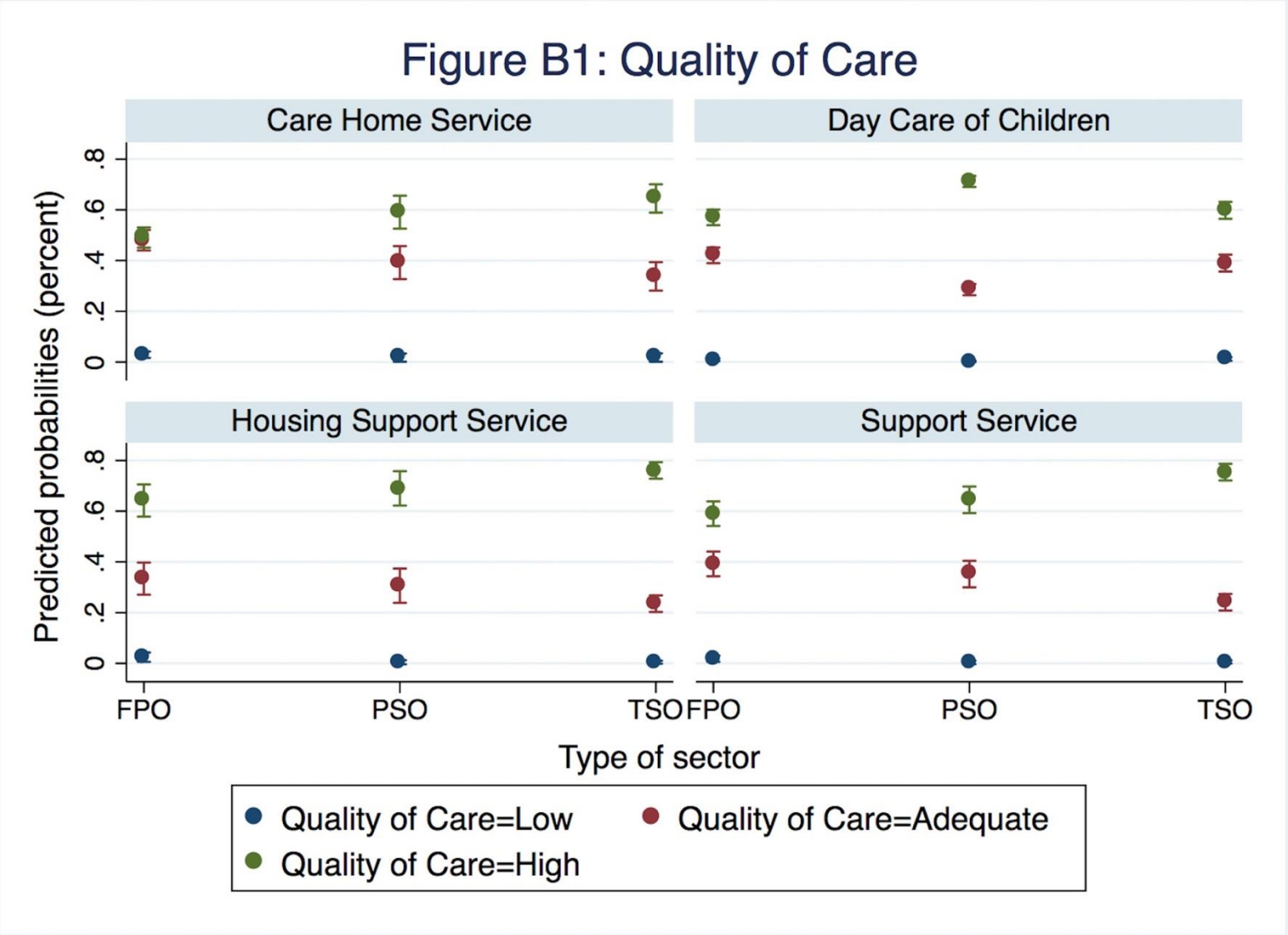
------------------------------------------------------------------------------
       complain_dum | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------------+---------------------------------------------------------
             ordsec |
                  1 |   .1827634   .0211099   -14.71   0.000     .1457376    .2291957
                  2 |   .2957996   .0295132   -12.21   0.000     .2432594    .3596876
                    |
                old |    2.79912   .6581929     4.38   0.000       1.7655    4.437879
                age |   1.000432   .0090009     0.05   0.962     .9829453     1.01823
         active_dum |   .9815351   .3415745    -0.05   0.957     .4962351    1.941441
         child_care |    4.37012   4.519341     1.43   0.154     .5757406     33.1711
        NumberStaff |   1.014777   .0013417    11.09   0.000     1.012151     1.01741
           children |   .4217847   .1519222    -2.40   0.017      .208208     .854445
housingsupport_dummy |    2.05948   2.260325     0.66   0.510     .2396322    17.69987
supportservice_dummy |   2.314599   2.539536     0.76   0.444     .2694963    19.87919
      carehome_dummy |   2.485428   2.666338     0.85   0.396     .3035585     20.3498
          school_dum |   6.472596   7.120815     1.70   0.090     .7492569    55.91474
            dementia |   1.207942   .2133104     1.07   0.285     .8545399    1.707496
                     |
        grade_change |
                  1  |   2.211787    .337599     5.20   0.000     1.639906    2.983098
                  2  |   2.590119   .3731485     6.61   0.000     1.952946    3.435176
                  3  |   .1189318    .033487    -7.56   0.000     .0684904    .2065219
                     |
               _cons |    .105057   .1212316    -1.95   0.051     .0109439    1.008505
------------------------------------------------------------------------------
```

Requirements

Logistic regression — Number of obs = 6320; LR chi2(16) = 755.05; Prob > chi2 = 0.0000; Log likelihood = -2920.0863; Pseudo R2 = 0.1145

requirement_dum	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ordsec						
1	.4935881	.0456649	-7.63	0.000	.4117321	.5917178
2	.7069104	.0608671	-4.03	0.000	.5971361	.8368649
old	2.628148	.5195157	4.89	0.000	1.783979	3.871774
age	.9730691	.0076567	-3.47	0.001	.9581774	.9881924
active_dum	1.113632	.3601256	0.33	0.739	.5908544	2.098953
child_care	1.228315	.6878283	0.37	0.713	.4098795	3.680977
NumberStaff	1.009016	.0012177	7.44	0.000	1.006632	1.011406
children	.8969939	.2154522	-0.45	0.651	.5601912	1.436292
housingsupport_dummy	1.487114	.9103079	0.65	0.517	.4480202	4.936177
supportservice_dummy	1.201242	.7388358	0.30	0.766	.3598244	4.010241
carehome_dummy	1.695957	.9899672	0.90	0.365	.5401942	5.324512
school_dum	1.13933	.7635617	0.19	0.846	.3063283	4.237522
dementia	1.229638	.2092639	1.21	0.224	.8808808	1.716476
grade_change						
1	1.335123	.1949419	1.98	0.048	1.002853	1.777484
2	3.674331	.4758097	10.05	0.000	2.8507	4.735929
3	.8678449	.1995073	-0.62	0.538	.5530457	1.361831
_cons	.2478778	.1725359	-2.00	0.045	.0633515	.9698805

Appendix C3: Predicted probabilities across individual service types



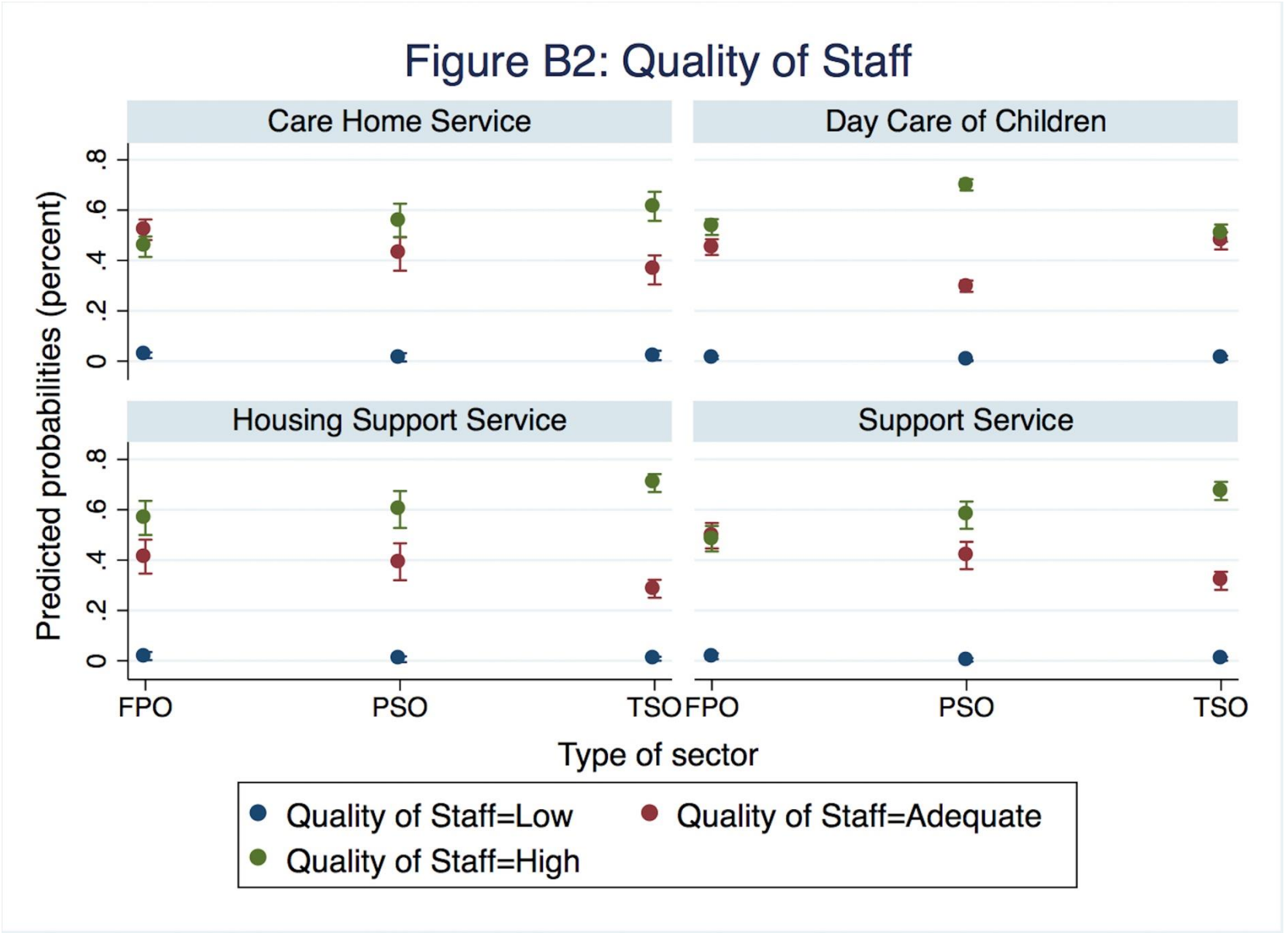

Figure B2: Quality of Staff
Care Home Service
Day Care of Children
Housing Support Service
Support Service
Predicted probabilities (percent)
0
.2
.4
.6
.8
FPO
PSO
TSO
Type of sector
Quality of Staff=Low
Quality of Staff=Adequate
Quality of Staff=High

Figure B3: Quality of Management

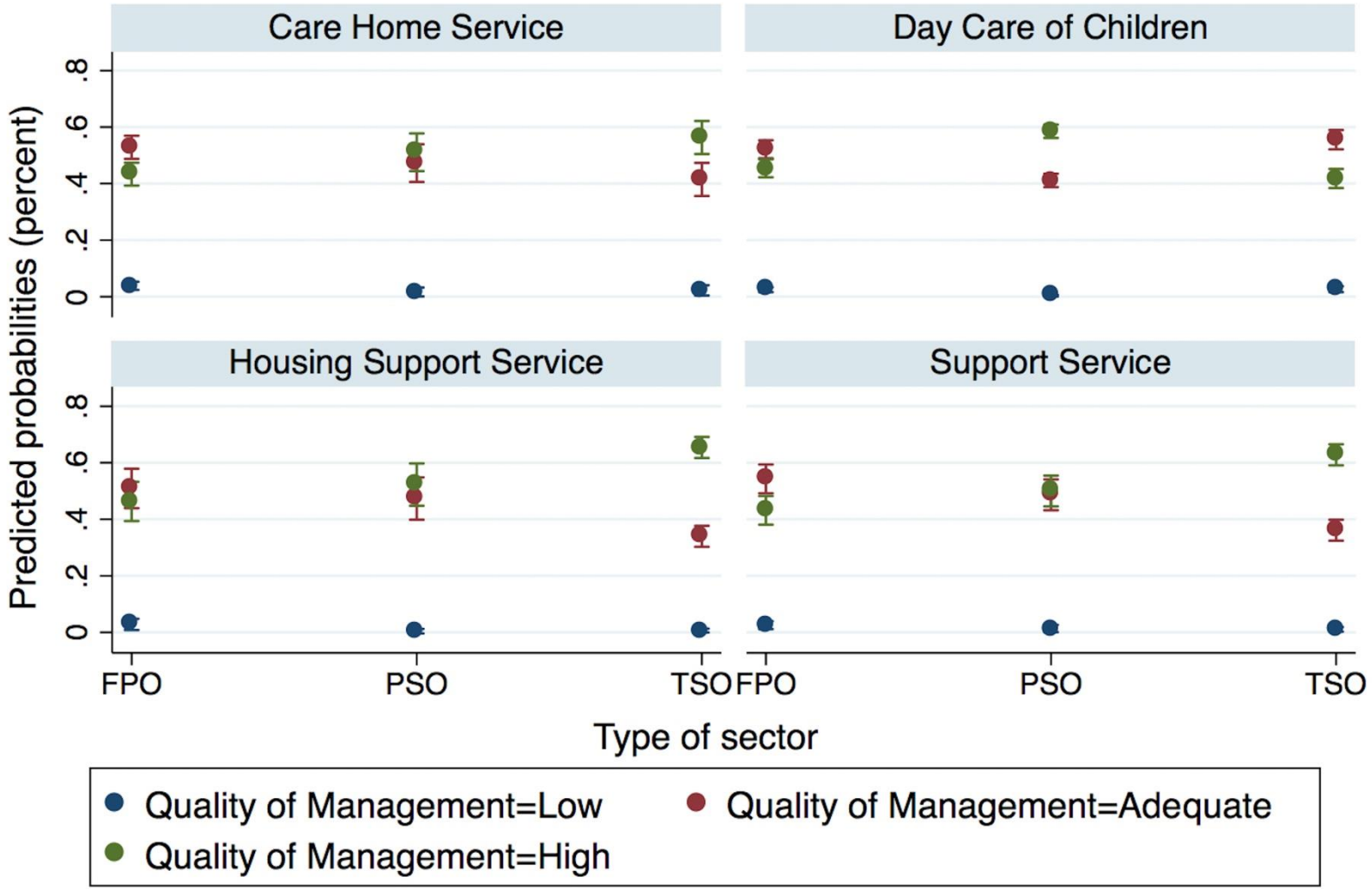
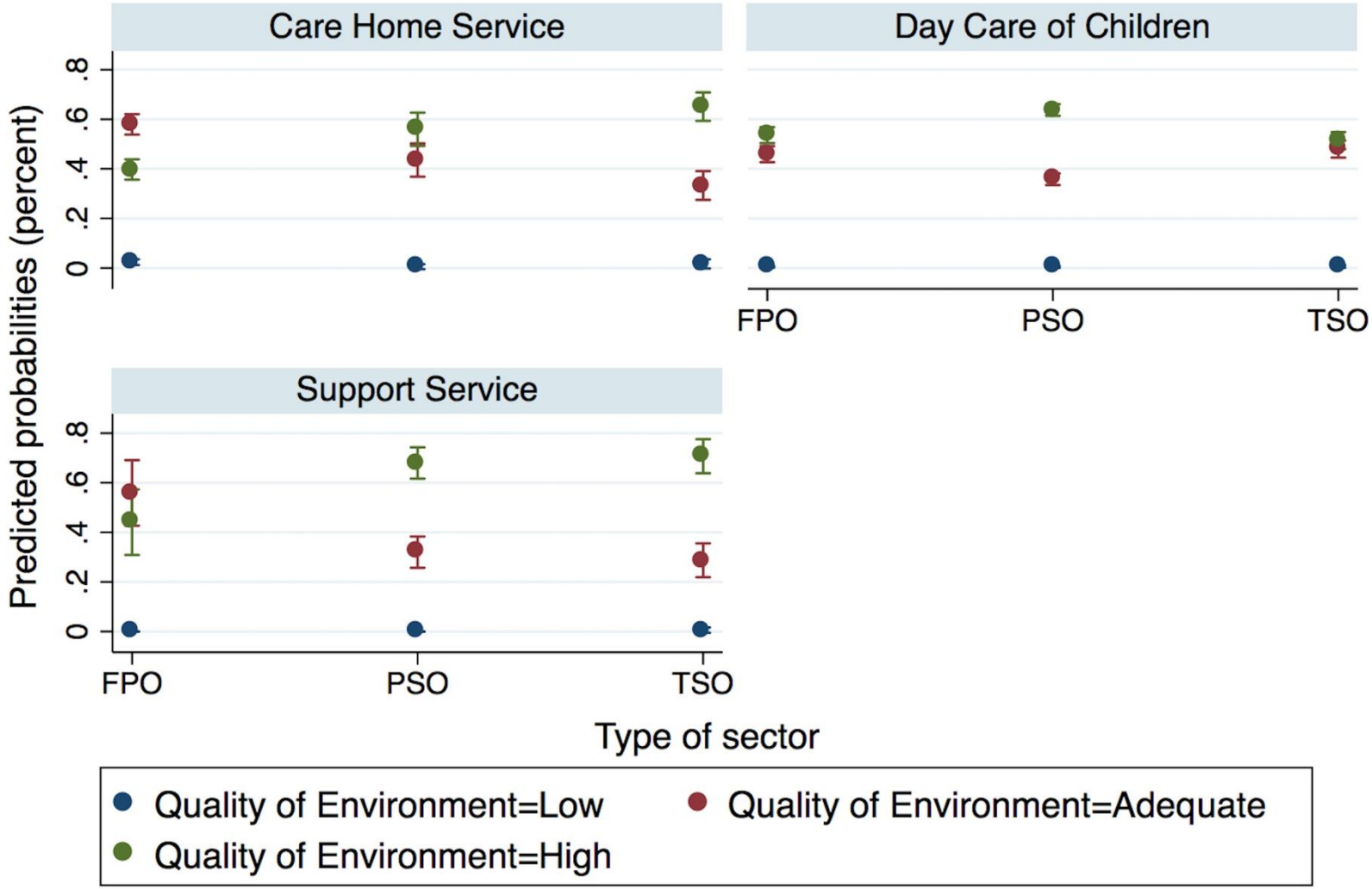
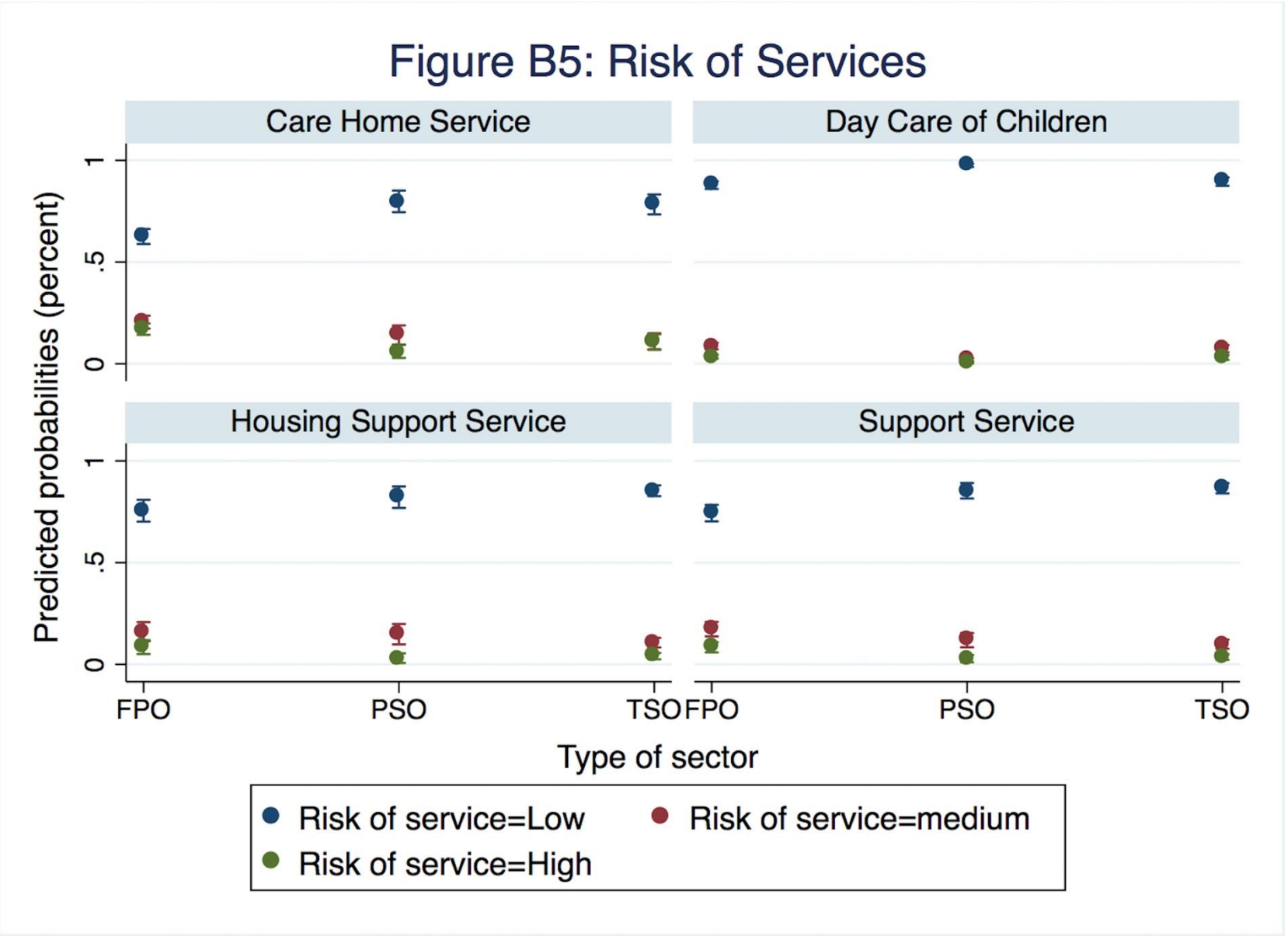


Figure B4: Quality of Environment




Figure B5: Risk of Services
Care Home Service
Day Care of Children
Housing Support Service
Support Service
Predicted probabilities (percent)
0
.5
1
FPO
PSO
TSO
Type of sector
Risk of service=Low
Risk of service=medium
Risk of service=High