Achieving Outcomes in Complex Public Service Systems: The Case of the Early Years Collaborative

A thesis submitted in fulfilment of the degree of:

Doctor of Philosophy

By

Max Nealon French

Faculty of Social Science

University of Stirling

Scotland

April 2017

Declaration

I declare that none of the work within this thesis has been submitted for any other degree at any university. I declare that all of the work within has been composed by the PhD candidate.

Acknowledgements

Abstract

Governments around the world have been increasingly adopting an 'outcomes-focus' in the design of policy and the management of public services, the implications of which have been subjected to increasing scrutiny within public administration (Boyne and Law 2005; Elvidge 2012; Heinrich 2002; Housden 2016; Lowe 2013; Lowe 2017; Wimbush 2011). Yet wherever an outcomes-based approach has been applied, be it within performance management (Bevan and Hood 2006; Lowe 2013; van Thiel and Leeuw 2002), budgeting (Perrin 2006; Ryan 2003), or commissioning, its achievements have fallen short of expectation (Wimbush 2011).

Outcomes have predominantly been conceptualised and operationalised within what this thesis calls a 'Rationalist' approach, linked to the New Public Management context within which an outcomes-focus was popularised. This approach assumes we can understand the factors which drive outcomes, plan appropriate service interventions, harness the resources and commitment needed to put such interventions into practice, and manage such interventions towards their expected end points. Outcomes however are inherently complex phenomena – they are always transboundary, always co-produced by the individuals who experience them, and always impacted by a large number of unpredictable and uncontrollable factors in their external environment. Public management theory and practice finds itself at a crossroads: an imperative to improve outcomes, and a paradigmatic inability to do so – a challenge which scholarship is just beginning to respond to (Housden 2016; Lowe et al. 2016).

This thesis contributes an alternative 'Complex Systems' theoretical framework which responds to (rather than simplifies or externalises) the inherent complexity which outcomes present. This theoretical framework draws on complex adaptive systems theory to enable a 'Complex Systems' approach to the management of outcomes. The framework is based on

the conception of outcomes as emergent products of complex systems, and integrates three defining components of complex adaptive systems (self-organisation, distributed agentic learning, and attractor states) to enable an endogenous process of service transformation in conditions of uncertainty. This theoretical framework provides public management with more solid footing for understanding, analysing and designing outcomes-focussed interventions, with distinct advantages relative to existing outcomes-based approaches, in pursuing complex public service outcomes.

The thesis applies this framework through a multiple embedded case study analysis (Yin 2009) of the Early Years Collaborative, a large-scale multi-agency Quality Improvement Collaborative operating across Scottish local authorities, as it seeks to improve a set of population-level child development outcomes. The empirical analysis makes three contributions.

Firstly, in road-testing the Complex Systems theoretical framework, the empirical analysis confirms its evaluative utility. The framework focuses analysis on three components of adaptive capacity: the capacity to learn, the capacity to self-organise based on learning, and the capacity to influence system behaviour through the manipulation of performance attractors. Taken as a whole, the theoretically-informed analysis shows that activating these capacities required significant deviation from the Quality Improvement Collaborative model on which the EYC was based. The ability to generate consistent explanations of performance in complex environments demonstrates the conceptual value of the framework as an interpretative tool.

Secondly, the theoretical framework is adapted to provide a novel framework for understanding how learning and improvement can be generated through service user co-production. This analysis provides rare micro-level empirical evidence which ties forms of co-production to discrete outcomes. This clarifies how service user feedback can improve the efficiency (inputs-outputs), effectiveness (outputs-outcomes) and relevance (redefining outcomes) of public services through differential pathways. In so doing, it contributes to a

growing interest in the creative potential of co-production for public service improvement (Bovaird et al. 2017; Bovaird and Loeffler 2016; Voorberg et al. 2014), and a drive within co-production research for greater clarity over the expected outcomes of different forms of co-production (Loeffler and Bovaird 2016; DuRose et al. 2017).

Thirdly, the analysis contributes to a growing international interest in the potential for improvement collaboratives as population-outcome focussed interventions (Bryk et al. 2011; Ghandour et al. 2017; Inkelas and McPherson 2015). This thesis contributes the first significant empirical study of a large-scale multi-agency improvement collaborative, and finds that the viability of model in a population-outcome context is challenged by three factors: an innately less measurable social environments which impairs the quantitative-focussed improvement methodology, the more significant fragmentation and poorer coordination in social service systems; and the more significant contextual differentials across which learning must transfer. The analysis concludes, in contrast to much of this body of scholarship, that improvement collaboratives must significantly alter their underlying methodology if they are to become viable in this extended role.

Table of contents

14

List of Tables

List of Figures

List of Acronyms

BSC	Breakthrough Series Collaborative
CPP	Community Planning Practice
EYC	Early Years Collaborative
GIRFEC	Getting it Right for Every Child
EYQIU	Early Years Quality Improvement Unit
MCQIC	Maternity and Children Improvement Collaborative
MFI	Model for Improvement
PDSA	Plan do study act
QIC	Quality Improvement Collaborative
RAFA	Raising Attainment for All
RO	Research Objective
RQ	Research Question
SPSP	Scottish Patient Safety Programme
TOC	Test of Change

Chapter 1. Introduction

Thesis background

The implications that public service outcomes hold focus for the design and delivery of public services has been a topic of increasing scrutiny within public management scholarship over the past 15 years (Boyne and Law 2005; Hienrich 2002; Lowe 2013; Lowe 2017; Lowe and Wilson 2015; Perrin 2006; Petch et al. 2013; Wimbush 2011). Such an 'outcomes-focus' has been a major driver of public service reforms internationally (Perrin 2006; Wimbush 2011), and within the UK and its devolved administrations in particular (Elvidge 2012; HM Government 2011; Housden 2014; Housden 2016; Law 2013). Manifestations of an outcomes approach can be seen in performance management approaches (Friedman 2005), commissioning strategies (Paley and Slasberg 2007), innovations such as Social Impact Bonds (McHugh et al. 2013), and benchmarking frameworks (McAteer and Stephens 2013). In Scotland, the implementation in 2007 of a whole-of-government focus on outcomes has been the lynchpin of a collaborative and people-centred approach to public sector reform (Christie Commission 2011; Housden 2014; Housden 2016).

Three reasons can be argued to have driven a focus on outcomes within public management scholarship and practice. Firstly, a normative drive particularly within social care and social policy maintains that outcomes, representing 'what matters' to the users of public services, should be the focus of strategic planning and performance management (Ball et al. 2004; Cook and Miller 2012; Nocon and Qureshi 1996; Qureshi and Nicolas 2001). Secondly, a technical-managerial drive, drawing from Public Choice theory and popularised within New Public Management reforms, has held that private incentives can be aligned with public interest by outcomes functioning as performance targets with clear lines of accountability (Boyne and Law 2006; Friedman 2005; HM Govt 2011; Schedler and Proeller 2010). Thirdly and most significantly for governments themselves is the growing realisation that the failure to improve or sustain outcome levels, combined with a 'permanent' austerity in the UK (Cameron 2010), constitutes a threat to the sustainability of future public services (Christie

Commission 2011; CIPFA 2013; Elvidge 2012; Housden 2016; Lowndes and McCaughie 2013).

For these three reasons, public service outcomes can be considered one of what Roberts (2014) terms the 'large forces' driving reform in public service systems, in way which extends the 'administrative limits' of public administration theory (Hood 1976; Hood 2010). Outcomes in this context pose a 'big picture' problem for public administration and management scholarship (Pollitt 2016): how can public service outcomes be sustained or improved in the face of increasing demand and stagnating levels of input?

The obvious response to this problem is to improve efficiency (input-output ratios), by sweating assets, improving employee performance, or reorienting resource from less to more productive areas. Confounding this logic however is an additional threat to public service sustainability: the failure of a generation of New Public Management reforms, focussed on competition (Le Grand 2009), 'targets and terror' (Coulson 2009) or 'naming and shaming' (Bevan and Wilson 2013) to improve outcomes on anything approaching a consistent basis, while often achieving the opposite (Hood and Dixon 2016; Pollitt 2013).

This thesis argues that outcomes have a number of unique characteristics which have not been consistently recognised by public administration scholarship. Outcomes are always co-produced by service users (and indeed, by non-users); they are transboundary problems and impacted by many different agencies and social forces; they are causally complex, characterised by a high degree of uncertainty and ambiguity; and they are fiscally recursive, both attracting public spending and impacting upon it in the future. For these reasons, this thesis argues that outcomes are better conceived as the emergent products of complex systems (Jayasinghe 2011; Lowe et al. 2016), rather than merely an extension of a public service production process (Pollitt and Bouckaert 2004; Boyne and Law 2006; Schedler and Proeller 2010).

However, dominant approaches to outcomes-based strategic planning and performance management – performance benchmarking, Social Impact Bonds, payment-by-results

schemes, results-based management and management-by-objectives, outcomes-based performance management – are rooted within a 'Rationalist' worldview, combining a Newtonian or scientific rationalist philosophical position with an economic rationalist view of human behaviour. Consistent with arguments that outcomes-based management represent an extension of the rational behavioural assumptions underpinning New Public Management and neoliberalism more generally (Smyth and Dow 1998), the lack of robust theoretical alternatives leave public managers and policymakers in a familiar position with respect to that ideology – adopt the Rationalist approach to outcomes management or abandon an outcomes-focus: there is no alternative.

Research logic and structure

The principal contribution of this thesis is to contribute conceptually, theoretically and empirically to emerging research across public management (Lowe et al. 2016; Lowe 2017), public health (Burns 2015), social epidemiology (Jayasinghe 2011), and collaborative service improvement (Bryk et al. 2011; Inkelas and McPherson 2015; McPherson et al. 2015) which have outlined complexity theory as a potential route beyond this impasse. The central Research Aim therefore is: *to enhance the ability of public service systems and interventions to improve outcomes by advancing theoretically and empirically a complex-systems based approach to outcomes management.*

This thesis' principal contribution to public management literature is to construct and operationalise an alternative 'Complex Systems Approach' to outcomes management which responds directly to the inherent limitations of the dominant 'Rationalist Approach' in tackling complex outcomes. This theoretical framework is applied to the Early Years Collaborative (EYC), a large-scale multi-agency Quality Improvement Collaborative (IHI 2003; Plsek 1999), as it operates to achieve four stretching population-level child development outcomes. While not explicitly or consciously adopting the framework developed here, the experience of the EYC as an outcomes-focussed and highly autonomous improvement network has strong potential to advance theoretical development in this area.

This thesis contributes to the Research Aim through the sequential achievement of Research Objectives (ROs) at the conceptual (RO1), theoretical (RO2) and empirical level (RO3), and through drawing implications for further research (RO4). This sequential structure guides the logic of this thesis' inquiry, however it also allows the insight developed through the literature review, theoretical, and empirical chapters to inform more specific and theoretically-relevant Research Questions (RQs) in response to important gaps in the literature. Three research questions are developed which both contribute to the research aim, and respond to contemporary debates in relevant scholarly literatures. The thesis structure of Research Aim, ROs and RQs provides a hierarchical logic which link together ambition, conceptual and theoretical development, and finally empirical data into a cohesive whole. These components are outlined below in Table 1.

Table 1. Research structure and logic

Aim	Objective	Associated Research Questions	Location of contribution
To enhance the ability of public service systems and interventions to improve outcomes by advancing theoretically and empirically a complex-systems based approach to outcomes management.	RO1 To develop a deeper understanding of the implications which outcomes hold for public management and the design of public services interventions		Chapter 2
	RO2 To theorise an alternative approach to outcomes management consistent with a complex systems view		Chapter 3
	RO3 To apply this theoretical framework to an empirical case	RQ1 How effective is the theoretical framework in an evaluative or interpretative capacity?	Chapters 5, 6,7, 8

	RO4 To draw implications from the research findings to advance public administration and social and public policy, and consider how such an agenda may be advanced in theory and practice.	RQ2 How can arrangements of co-production lead to improved service outcomes, and what implications does this hold for a complex systems approach to outcomes management?	Chapter 7
		RQ3 Can the Quality Improvement Collaborative be an effective approach to outcomes-focussed improvement?	Chapter 8
			Chapter 9

Route map to the thesis

The structure of this thesis can be considered as two parts. In the first part, the thesis' main conceptual (Chapter 2) and theoretical contributions (Chapter 3) are made, and the study is primed for empirical analysis (Chapter 4). In the second part, the empirical data is reported and interpreted and the three RQs are tackled. Chapters 5 and 6 report the empirical data, while the Complex Systems theoretical framework is applied in Chapters 7 and 8. Chapter 9 finally takes stock of thesis' findings and considers the implications for public management theory and practice. The following section concludes the thesis introduction by providing the reader with a more detailed route map to the thesis, which locates the specific contributions made against its RQs and ROs.

Chapter 2 undertakes a critical literature review of the role of outcomes in governance and public management. It finds that a dominant 'Rationalist' approach to outcomes management, combining the philosophical rationalism of Newton and Descartes with an economic rationalism underpinned by Public Choice theory, has manifested in approaches such as performance benchmarking, league tables, results-based management and payment-by-results schemes. Challenging this view however is an emerging Complex Systems approach, drawing from social epidemiology and public health literature, which views outcomes as the emergent product of complex systems and takes a worldview aligned with critical realism. This latter approach remains theoretically underdeveloped however, despite implicit recognition in the function of outcomes as 'shared goals' in multi-agency partnerships and as measures of 'individual value' in personalised service approaches across health and social care.

Taking this conceptual development as its starting point, Chapter 3 begins a concise theoretical exploration of Complex Adaptive Systems theory to consider how a complexity-consistent alternative, here called the 'Complex Systems' approach might be put into practice. It integrates three fundamental elements of Complex Adaptive

Systems – self-organisation, distributed learning and attractor states – to develop a theoretical framework for understanding change and transformation in public service systems consistent with the view of outcomes as emergent. The applicability of the resulting 'Complex Systems theoretical framework' is then discussed in reference to the subject of the empirical research, the Early Years Collaborative.

Chapter 4 then sets out the research design for the empirical analysis. It begins with a description of the thesis logic and structure, describing in greater depth the origin and significance of the study RQs. It situates a complexity worldview within a critical realist philosophical paradigm, describes and justifies the case study methodology employed. The multiple embedded case study methodology employed (consisting of three case study areas and eleven in-depth units of analysis), the purposive sampling method, the data collection and analysis procedures and research process are then described and justified.

Chapter 5 begins the empirical analysis with a descriptive and comparative account of the development and strategic trajectory of the EYC both nationally and within case study areas. The comparative analysis surfaces a number of key similarities and differences in strategic and operational management across the case study areas. It finds little evidence that the EYC functioned as a genuine learning system, with improvements and strategic focus remaining localised.

Chapter 6 completes the exposition of empirical data with the presentation and detailed empirical analysis of the 11 embedded case studies of improvement projects. This chapter critically examines the improvement process within the EYC as it operates across different service contexts in the planning, testing and scaling of improvement. It finds that key methodological difficulties inhibited learning and innovation, which implies for RQ3 that QICs operating at a population-level context are likely to face many additional barriers.

Drawing on the full range of data presented in Chapters 5 and 6, Chapters 7 and 8 begin more focussed and direct enquiries into thesis RQs. Chapter 7 applies the theoretical framework to the embedded case studies, undertaking a more forensic analysis of learning and its adaptive effects. Drawing on a comparative analysis of the Locality Model (an overlapping initiative with the EYC in one case study area), the framework integrates co-production with adaptive learning using the theoretical framework, and situates findings within a burgeoning literature on public service co-production (Osborne 2016), answering RQ2.

Chapter 8 applies the theoretical framework in an interpretive capacity to the empirical data introduced in Chapters 5-7, to consider how the EYC demonstrated learning, self-organisation and coordinative capacity. The application of this model fulfils RO3, and permits a deeper theoretically-informed analysis of the Quality Improvement Collaborative model on which the Early Years Collaborative is based, answering RQ3. The application of the framework isolates three particular barriers which must be surmounted by QICs operating to improve population-level outcomes.

Finally, Chapter 9 summarises the thesis' contributions to public management theory and practice, and draws insight developed across the chapters towards a summative contribution to the Research Aim. It uses this as a springboard to tackle RO 4, considering how further research and practice-oriented inquiry can advance outcomes-based approaches in a way which transcends the limitations of the Rationalist approach.

Chapter 2. The implications of an outcomes-focus for public management

This chapter undertakes a critical literature review of the role of outcomes in public services and its implications for public management and service reform. The literature review opens with a historical review of the use of outcomes in public management, which reveals significant differences in the roles outcomes have come to play over the past fifteen years. The review moves on to consider the dominant 'Rationalist' approach to outcomes management, which views outcomes as results which emerge through an intra-organisational production process. It then discusses an alternative 'Complex Systems' conception, based on a social-epidemiological model, which envisages outcomes as the emergent product of complex service systems. These two approaches are compared on the ontological, epistemological and theoretical levels. The chapter concludes by arguing that while a Complex Systems approach appears conceptually superior, it remains methodologically and theoretically underdeveloped. This chapter seeks to meet Research Objective 1: *to develop a deeper understanding of the implications which outcomes hold for public management and the design of public services interventions.*

Outcomes and public management: a historical review

Outcomes – defined in the broadest sense as indicators of societal progress – have been an important philosophical and political issue since at least the late 18th century. The Scottish Enlightenment figure John Sinclair, more famously credited with the popularisation of the term 'statistics', oversaw the *Statistical Account of Scotland* in 1791, the first robust national account of social, economic and agricultural conditions, comprising a 160-question survey delivered across over 900 parishes. Sinclair described the intention behind the Account as 'an inquiry into the state of a country, for the purpose of ascertaining the quantum of happiness enjoyed by its inhabitants' (Sinclair 1798, p. xiii). Prior to this, national statistical accounts had 'uniformly been

instituted, with a view of ascertaining the state of the country, for the purposes of taxation and war, and not of national improvement' (Sinclair 1797, p.xxxv). Sinclair's intention, in line with the empiricist philosophy underpinning the Enlightenment, was to understand social issues within populations which could guide national efforts for social improvement.

This ambition was taken up notably by the Statistical Societies established in major English and Welsh cities across the 1830s. The Manchester and London Statistical Societies – the only two still extant – both began statistical analysis of aggregate measures of social problems including crime, disease, and urban squalor. Social outcomes measured through statistics became 'the "empirical arm" of political economy' (Porter 1986, p.27), used for instance by William Godwin (1820) in his *Of Population* to rebut Thomas Malthus' gloomy predictions of catastrophe through unchecked population growth and the necessity of poverty as a limiting force.

Outcomes also began to take hold during this period within the wider development of the social sciences. Engels' (1845) analysis of poor living conditions and disease amongst the English working class cemented a structural link between social conditions and the clustered nature of health outcomes. Emile Durkheim's (1897) foundational sociological studies of suicide rates influenced understandings of suicide as both a social and an individual behavioural phenomenon and set the stage for a sociological view of structure and agency (Parsons 1949). Sociological engagement with societal outcomes was thus rooted in a view of outcomes as systemic properties structured by social problems.

John Sinclair's concern with indicators of social progress found particular impetus in the post-war twentieth century, as Bovaird (2014) notes, particularly within public health and quality of life. Zeckhauser and Shepard's (1976) development of the QUALY (quality-adjusted life years) for instance, incorporated wellbeing and enjoyment

alongside hard measures of mortality to aid clinical decision making, bringing social concerns further into epidemiological thought.

International organisations have also been at the forefront of using cross-national comparisons of social progress to inform better policymaking. The UN's Human Development Reports introduced in 1990, echoing John Sinclair's emphasis on national wellbeing improvement, aimed to 'to shift the focus of development economics from national income accounting to people-centred policies' (Sen 2000, p.20). The OECD's Society at a Glance publication initiated in 2001 has listed a growing a range of cross-national social indicators for policy improvement, while the OECD Better Life Index established in 2011 on recommendation of the Stiglitz-Sen-Fitoussi Commission (Fitoussi et al. 2009) is the first major multinational attempt to establish comparable indicators of well-being.

Outcomes as performance targets within New Public Management

Given this historical context it is notable that outcomes lacked any comparative influence on theory development in public administration for much of the 20th century. Public administration emerged as a coherent discipline following Woodrow Wilson's dichotomy of politics and administration (Wilson 1887). The period of 1945-79, characterised by the rise of the post-war welfare state and large public bureaucracies was influenced by Weber's (1922) idea of a 'perfect bureaucracy', which could uphold values of universalism and fairness in the administration of public goods. Public management theory during this time was strongly influenced by mechanistic approaches to process management, such as Supply Chain Management (Bovaird 2014). While outcomes might continue to function as background indicators of social progress and so affect policy development, they had no operational relevance to managerial or administrative decision-making, as they had within epidemiology, policy analysis or sociology. Simon (1957, p.xxxvi cited in Bovaird 2014) for instance argued in such a context that 'high level goals provide little guide for action'.

Criticism of this 'traditional' public management theory emerged from proponents of what came to be known as the New Public Management (NPM) (Dunleavy and Hood 1994; Hood 1991). Osborne and Gaebler's (1992) influential call for reform *Reinventing Government* warned that lacking signals of effectiveness from a market mechanism, public service officials were free to pursue their own private interests at the public's expense. Advocates of NPM reforms advised moving from a focus on bureaucratic compliance and implementation, to a focus on the *results* of public intervention, and accordingly the management of service outcomes and outputs rather than inputs.

NPM was strongly influenced by the laissez-faire Austrian School of economics which asserted the centrality of economic self-interest to economic progress, and the Public Choice theory was extended the same assumptions to the public realm (Buchanan and Tulloch 1962; Olson 1965; Ostrom and Ostrom 1971; Niskanen 1975). Public choice theory was built around the behavioural assumptions of John Nash's non-cooperative game theory, which assumed total self-interest, and with it the predictability of collective behaviour. Public choice theory transferred this economic rationality to public servants, whose rational self-interest amongst undermined any 'public service' duty to delivering public goods. The solution, building on Peter Drucker's (1954) Management by Objectives, was to set performance targets to align private incentives with the public interest. Outcomes were first operationalised within public management as the means to accomplish this task.

Where competition between providers was not possible politically or practically, free market think tanks, most prominently the Adam Smith Institute and the Institute for Economic Affairs, advocated the use of outcomes as shared performance 'targets' against which the performance of public service staff could be incentivised (James 1993). Extrinsic incentives would reward 'good performance' while sanctions would be meted out for failure or underperformance. Such regimes could theoretically be used to drive efficiency, since poor performers could be weeded out, and workers could be

motivated towards better performance. Outcomes, which previously functioned only as background indicators, became an active driver of efficiency.

Targets continued to be advocated by John Major's Conservative government and were embraced even more enthusiastically under New Labour's 'modernisation' agenda (Cabinet Office 1999). In a speech following New Labour's election in 1997, Prime Minister Tony Blair announced his intention to lead 'a government that focuses on the outcomes it wants to achieve, devolves responsibility to those who can achieve those outcomes and then intervenes in inverse proportion to success' (Blair 1998, p.63). Outcomes continued to be synonymous with performance targets, and came to characterise the relationship between local and central government. Local governments were expected to work towards targets imposed by central government under the Best Value regime, and later through Public Service Agreements.

Around this time, outcomes also became operationalised as integrated management tools for results-based management. In particular, Mark Friedman's (2005) Outcomes-based Accountability (known as Results-based Accountability in the US) has been extremely influential in shaping discourse around outcomes (Mayne 2007). While carrying through the conditions of measurability and performance incentives central to results-based management, Outcomes-based Accountability emphasises that target-driven behaviour requires allowing practitioners the freedom to experiment and respond creatively to improving practice.

The conditions of measurability, 'carrot and stick' performance incentives, and accountability to measured performance against pre-defined targets have been integral features of new approaches to outcomes-based commissioning. Most prominently, these include Payment by Results approaches, in which payment is stipulated on an agreed level of measurable achievement against pre-defined results (HM Govt 2011), and Social Impact Bonds, a financial investment model based on Payment by Results, but which attracts private investment and capital market

involvement (Social Investment Task Force 2010). Payment-by-Results approaches which have been used in employment, reduction of criminal recidivism, services for those deemed 'troubled' families, NHS commissioning and international aid (Morse 2015). These models are intended to encourage a process of competitive innovation for contracts, echoing Friedman's (2005) call for a freedom from bureaucracy in the pursuit of better results.

The literature review so far has documented the genesis of outcomes as a statistical measure of social process towards an operational component of public service management within NPM reforms. However, over the last fifteen years, outcomes have seen significant diversification in their conceptualisation and usage within the management and administration of public services. There have been two notable areas of development which have become particularly widespread, and will be discussed here: shared goals and personal outcomes.

Outcomes as shared goals

NPM reforms resulted in a fragmentation of management and service delivery functions and 'hollowing' of the state, which heralded the subsequent rise of 'networks' of governance, service delivery (Kickert et al. 1997; Rhodes 1997). Network responses are justified for their necessity in tackling cross-boundary problems – particularly of the deep-seated, ill-defined and intractable order which Rittel and Webber (1973) describe as 'wicked problems' (Ferlie et al. 2011; Roberts 2000) – through bringing the required resources and expertise to tackle cross-cutting issues (Sullivan and Skelcher 2002). The proliferation of such networks across service delivery, policy formation, and governance over the past 20 years (Lecy et al. 2014) have frequently entailed the use of cross-boundary outcomes as shared goals.

This role was prominent in comprehensive area-based initiatives such as New Labour's New Deal for Communities and Health Action Zones, in which articulated population-level outcome indicators were used to inform action and assess performance (Ling

2002). Outcomes played a similar role in the later place-based 'Total Place' pilots (Bovaird 2010), and within broad-based collaborative endeavours, such as the subject of this thesis, the Early Years Collaborative (Scottish Government 2014).

In a recent review of outcomes approaches internationally, Wimbush (2011) finds the use of outcomes as shared goals in governance partnership settings to be the most significant contemporary development. The move to an outcome-based framework for local government was first trialled in 2001 in England under Local Public Service Agreements, which held local authorities to account based on an array of input, output and outcome measures. These were later replaced by 'Local Area Agreements' (LAAs) – suites of measures, including a greater share of outcomes, which governing multi-agency Local Strategic Partnerships in England were supposed to jointly tackle, with similar developments in Wales (Law 2013).

In Scotland, Community Planning Partnerships (CPPs) – comprising local authority and principal public agency partners – were established in 2003, and in 2007, a 'concordat' agreement struck between the minority Scottish National Party government and the Coalition of Scottish Local Authorities (COSLA) set out a 'Single Outcome Agreement' (SOA), a suite of outcomes agreed by local and central government as the basis for local service planning. Local government was afforded significant autonomy over which outcomes to pursue and how to go about this pursuit (Cairney 2015).

Outcomes have also increasingly found a role for outcomes as a focal point for budgeting and coordinating public resource across an increasingly fragmented governance landscape (CIPFA 2013; Perrin 2006). Governments in Australia have since 1999 required departments to produce an 'outcomes statement' linking plans in a transparent way to the achievement of a wider governmental purpose with New Zealand following suit shortly after (Ryan 2003).

The idea of outcomes-based government has taken hold in Scotland, with the Scottish Government adopting a model used by the Virginia State government (Virginia

Performs 2016) in 2007. While target-based incentive systems have been adopted (for example with Public Service Agreements), 'shared goals' have tended instead to function as a rallying point to motivate distinct actors around a shared purpose, and a reference point to motivate collaborative planning, monitoring and management activities. Scotland has been noted for the 'continued primacy' of outcomes, 'an understanding that this requires the head-to-toe alignment and integration of all involved in public services' (Housden 2014, p.64). This has been pursued through a collaborative and integrative programme of service reform epitomised by the Christie Commission (Christie Commission 2011), which has continued through three Permanent Secretaries (Elvidge 2011; Housden 2016). This approach has been noted internationally (Elvidge 2012), and was influential in Northern Ireland's decision to follow suit more recently (Northern Ireland Assembly 2016).

The role of outcomes as shared goals in this sense is typified by perhaps the best known set of outcomes internationally, the UN's Millennium Development Goals, and the Single Development Goals which replaced them in 2015. These goals require consensual multi-national responses to shared global problems where there is no real possibility of direct coercion or formal sanction structure (Hulme 2010).

Outcomes as individualised value

While outcomes are most often understood as aggregate statistical measures, it is increasingly been recognised that they are achieved through highly individual pathways. Scholarship emerging from the Social Policy Research Unit at the University of York since 1995 has positioned outcomes as measures of individual value, recognising that variation in user needs, preferences, values and particularities demand more responsive services and a greater voice for users in directing services (Novac and Qureshi 1996; Ball et al. 2004; Glendinning et al. 2009; Qureshi and Nicolas 2001; Qureshi et al. 1998). This literature implicitly responds to cybernetician Ross

Ashby's (1956) Law of Requisite Variety: that a viable system is capable of at least as much variety as that presented to it by its environment.

Under the 'personal outcomes' approaches which followed, such Personal Budgets and Self-Directed Support (Carr 2011), service users and their care networks are recognised as possessing vital experiential knowledge (Beresford and Boxall 2013), with a key role to play in the design of service responses. Personal outcomes approaches shift focus from 'matching needs to services' towards 'identifying what matters to the person and working out what role each person will have in achieving that outcome' (Miller 2010, p.10), resonating with Needham and Carr's (2009) discussion of co-production, which they argue involves actively negotiating service improvements between users and providers.

As measures of individual value, outcomes have created particular impetus within healthcare to develop patient-reported measures of outcome in health care as an alternative to aggregate indicators of material change, which outcomes functioning as targets and shared goals invariably represent (e.g. Clancy & Eisenberg 1998; Krumholtz 2008). These 'personal outcomes' approaches have reached a high profile in Scotland in no small part through the 'Talking Points' approach developed by Emma Miller and Ailsa Cook in partnership with the Scottish Government-funded Joint Improvement Team (Cook and Miller 2012). Talking Points seeks better outcomes through a 'conversational' approach with service users and carers and carries an integrated user-defined measurement framework (Miller 2010). The recognition that outcomes are individuated and achieved through diverse pathways has also underpinned the wider 'personalisation' agenda in health and social care, with innovations such as Self-Directed Support and personal care budgets seeing widespread implementation across the UK and Scotland (Audit Scotland 2014).

Comparing the three functionalities: what are outcomes?

This literature review has found three distinct functionalities through which outcomes have been operationalised, originating from different bodies of scholarship: outcomes as results in the context of NPM reforms, shared goals in the context of partnership working and network governance, and personal outcomes in the context of health and social care and person-centred public services. The principal characteristics of each functionality including their mechanisms of service improvement and examples of their application in the UK are presented in Table 2.

Table 2. Outcome functionalities within public management

Role	Scholarly background	Implications for public services	Incentive Mechanisms	Mechanism of improvement	Prominent examples in practice
Results	Public choice theory, New Public Management	Rationalisation of internal processes and innovation based on external goals	Extrinsic rewards encourage efficiency and facilitate competition	Competition	NHS 4-hour waiting targets; School league tables; Social Impact Bonds, Outcomes-based accountability
Shared Goals	Network governance, meta-governance	Joint working through forms of collaboration and integration	Shared goals provide a common goal and call to action	Collaboration	Single Outcome Agreements, area-based initiatives, improvement collaboratives
Individual value	Health and social care	Co-prioritisation of outcomes and co-design of the service response	Intrinsic motivation of staff to achieve what matters to their service users	Co-production	Talking Points; Self-directed support; Personal Care Budgets

In their role as results, outcomes have facilitated competition-based and performance management reforms with an ambition of improving efficiency. As shared goals, outcomes have responded to the public service fragmentation and the growing footprint of networks in governance and service delivery, facilitating joint-working and integration towards trans-boundary goals. As measures of individual value, outcomes have provided a way to implement person-centred services in health and social care, and provided an appropriate management framework for the delivery of highly individualised services such as social care packages.

Table 2 also illustrates an important difference in how outcomes are conceptualised across these three roles. As results, outcomes are the end point of a linear production process, and thus are achieved directly by service organisations through a rationalisation of internal processes. This creates the potential for a competitive mechanism to reduce costs and improve value. In the other two approaches, outcomes are viewed as entirely detached from public services, echoing their conceptual development as social indicators (Sinclair 1797) and sociological objects (Durkheim 1897).

As shared goals, outcomes are viewed as external transboundary problems which necessitate joint working and collaboration. This responds directly to the multi-sectoral and systemic nature of outcomes (Housden 2014): health and educational outcomes are not 'produced' by the internal processes of hospitals and schools, but by a much broader array of social determinants (Marmot 2005; Marmot et al. 2008). As individual value, outcomes are co-created and negotiated by users and providers (Needham and Carr 2009). This recognises that users as experts in their own care are vital inputs into the creation of effective service responses. It also recognises that unlike service outputs, outcomes are always co-produced by service users: a school produces no educational value if its pupils disengage with learning.

It is important to note these functionalities are not mutually exclusive. Targets can be integrated with shared goals under joint accountability regimes (Jones and Stewart 2009), while aggregate high-level shared outcome measures may still contain scope for adequate personalisation. However, they tend to exist in isolation and conditions of mutual exclusivity: thus shared goals are formed in networked situations where it is not possible or feasible to enact vertical accountabilities, and 'personal outcome' approaches tend to be practiced in health and social care situations where needs and preferences are highly personalised.

With the exception of 'targets', we can detect different motivations and philosophical assumptions within outcome functionalities. For example while the Scottish Government has articulated a whole-of-government approach to outcomes as an attempt to integrate service provision and reverse the hollowing of the state, the Social Return on Investment Working Group (2011, p.9) valued the potential of outcomes-based government 'to open up delivery of government commissioned activity to companies and the voluntary sector'. Thus an analysis of the role of outcomes within public management must go beyond their function within public services, and explore the deeper motivations, rationales and assumptions behind their adoption.

The Rationalist Approach to outcomes management

The previous section has described the three principal functionalities of outcomes in a public management context, however the chapter has not yet engaged with the underlying conceptual differences within which these functionalities are situated. This section discusses the dominant paradigmatic approach to the management of outcomes, which this thesis terms the 'Rationalist' approach.

As Boyne and Law (2005, p.253-254) argue, 'most discussions of performance measurement are based on an implicit model of the production process in the public sector'. Outcomes are most commonly understood within public management and evaluation literatures as occurring following these 'production processes' or 'value chains' (Pollitt and Bouckaert 2004; Boyne and Law 2005) which link inputs to outcomes through often complicated causal logic models (Bovaird 2014). This reflects a mechanistic model of service production whereby inputs are structured into outputs through logical pathways which model processes in public services (e.g. Ostrom et al. 1978). These outputs then interact with their environment and impact upon outcomes, following the simplified logic model adapted from Schedler and Proeller (2010) outlined below in Figure 1.

Figure 1. The Rationalist approach to outcomes management

Inputs
- Public resources (labour, goods, capital or services) are structured into service activities

Outputs
- These activities in combination result in a range of delivered outputs

Outcomes
- These outputs interact with their environments to trigger outcomes for service recipients
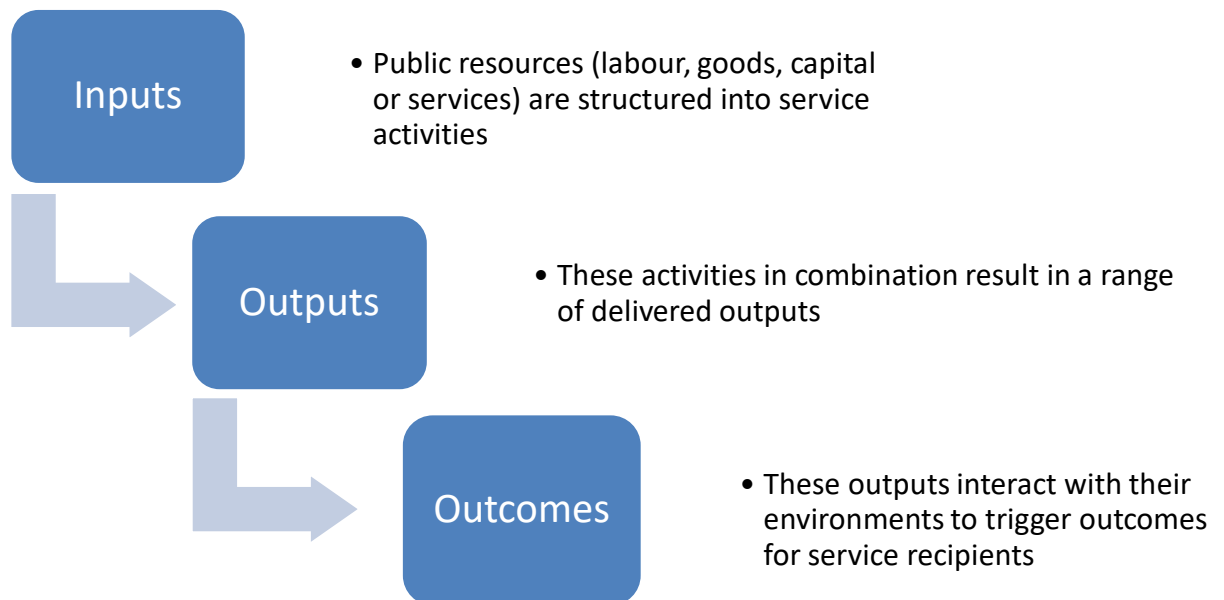
In Figure 1 'inputs' denote resource contributions (budget, staff and other resources) which are transformed through service processes into outputs (that which is produced – labour hours, tasks completed). Outcomes themselves are commonly broken down in programme theory into sequential stages, for instance 'proximal' and 'distal', or immediate, intermediate and final outcomes. Outcomes are distinct from outputs in that they exist beyond organisational boundaries, yet in this perspective, they remain at least partial products of the production chain. This understanding is typified by the conflation of many public management scholars of outcomes with 'results' or 'impacts' of service processes (Glendinning 2006; Pollitt and Bouckaert 2004; Schedler and Proeller 2010).

In order for attribution to be ascertained, outcomes must be measurable. Smyth and Dow (1998, p.295) for instance call outcomes 'singular, measurable, standardised and unequivocal second-order consequences'. Cartwright et al. (2016) describe how outcomes can be 'pinpoint' concepts – relatively precise and unambiguous measures like mortality or disease incidence – or can represent multifaceted and contested

concepts, such as poverty, disability or health. This latter group Cartwright et al. (2016) term 'Ballung' concepts, where measures can only imperfectly represent a wider concept, and which characterises perhaps the majority of meaningful social outcomes. Public management scholarship has struggled with the inherent complexity this measurement difficulty presents. The Audit Commission (2000) for instance published guidelines for outcome measurement, while Heinrich (2002) also provides pragmatic advice for selecting outcome indicators. Boyne and Law (2005) concede that many 'wicked issues' arise when translating 'Ballung' outcomes into quantitative indicators for performance management, including alignment between indicator and outcome, linkage between outcomes and organisational performance, and the specification of indicators which are not easily gamed.

If these technical measurement problems can be resolved, outcome indicators can be tracked, and crucially, actors can be held accountable to variation observed. This often involves some form of causal modelling which links inputs to outcomes based on an understanding of the problem system which interventions are designed to interact with (Bovaird 2014), often through an appeal to an existing evidence base. This process can also generate definitive knowledge about 'what works' as movement in outcome indicators becomes assigned to various processes in the production chain. This process has the conceptual benefit of allowing the calculation of programme effectiveness through value for money (outputs divided by inputs) and effectiveness (outcomes divided by inputs) through techniques such as Social Return on Investment (Scholten 2006).

Philosophical rationalism

Jayasinghe (2011) writes that public health has similarly adopted a 'Newtonian' conception of reality, citing as its basis Newton's *Principia Mathematica* which outlined the fundamental 'natural laws' of classical mechanics which described an objective reality, and Descartes' systematic inquiry which advocated the reductive

inquiry into complex systems. In public administration, these assumptions have been central to the production process model described in the previous section. Jayasinghe (2011) argues this approach can be summarised through three concepts:

- Linearity: inputs are proportionally related to outcomes.

- Reductionism: a larger causal system determining outcomes can be broken down into smaller constituent parts without compromising the integrity of the whole.

- Hierarchy: there is an ability to control the strategic orientation (through hierarchy or competition) of the actors needed to embark on a particular course of action.

These ontological assumptions accord with a scientific realist position that social phenomena are driven by verifiable and immutable natural laws. The stability and predictability afforded by conditions of linearity, reductionism and hierarchy permit experimental designs and statistical methods to yield objective knowledge about 'what works' in relation to outcomes.

Smyth and Dow (1998, p.291) writing about education argue that outcomes represent a drive to 'technologise schools, teaching and learning' and as part of a 'technocratic rationalism', frees decision making from 'the reliance on the teacher's value-laden, unreliable and subjective assessments' (Smyth and Dow 1998, p.298). In a manner redolent of Taylor's (1914) Scientific Management, Blair's Modernisation agenda, was based on the assumption that it was possible to determine objectively 'what works' in the policies and services that government commissioned. This technocratic approach to outcomes management assumes conditions of stability and predictability in line with the Newtonian or Cartesian worldview.

Economic rationalism

As discussed, outcomes first arose in an operational capacity in the context of late NPM reforms in the early 2000s, and were strongly influenced by Public Choice theory which assumed rational self-interest. Reducing outcomes to measurable targets allowed the reconciliation of the public interest with the private interest of officials within lower administrative units – be they local authorities, departments of government, teams or individual staff members, or institutions like schools and hospitals.

The exact means of this varied, for instance targets could be used for process optimisation under a Taylorist scientific management regime under overarching outcome goals, and through rational strategic planning relying on causal modelling (Bovaird 2014). Alternatively, such micro-management could be eschewed, for example through specifying monetary rewards and sanctions in outcomes-based performance management systems (Lowe 2013) enforcing 'targets and terror' (Coulson 2009) in the management of employee performance. In practice however the Rationalist Approach has increasingly tended to abstain from 'detailed' managerialism (Klijn 2008), instead externalising responsibility for outcomes through contractualism using Social Impact Bonds, Outcomes-based Commissioning, Payment by Results or Outcomes-based Accountability, or through benchmarking and league tables (Bevan and Wilson 2013).

Criticism of the rationalist approach

Smyth and Dow (1998, p.298) note that, 'while outcomes rhetoric may be value free, it is not value neutral'. The conflation within much public management literature of outcomes with 'results' reflects a tendency to frame debates about outcomes within a technocratic and methodological discourse, as opposed to concerns of epistemology, values or politics. The body of criticism which the rationalist approach has accumulated is substantial, as this section will summarise. It is helpful to distinguish between a set of epistemic limitations (knowing what to do), and a set of control

limitations (doing what we know), which have been extensively documented in the public management and evaluation literatures. These comprise:

- (1) The attribution problem – any number of factors may impact upon outcomes and it can be difficult to isolate the contributions made by a particular organisation or initiative (Bovaird 2014). This is particularly problematic where payment is linked to attribution (Lowe 2013; Lowe and Wilson 2015).

- (2) Unclear direction of travel from outputs to outcomes – outcomes which are long-term or distal, which can depend on many intermediate steps (Tunstill and Blewitt 2013), each of which may be uncertain or be poorly evidenced.

- (3) Poor understanding of causality – outcomes are driven by complex webs of causation which may be poorly understood (Bovaird 2014; (Ryan 2003), and have few evidence-based interventions at hand.

- (4) Time lags in impact – many health and social interventions underpinned by the move to a prevention focus incur upfront costs but are predicated on more long-term outcomes being achieved – or on negative outcomes being prevented (Boyne and Law 2005). Such 'maintenance outcomes' (Nicolas et al. 2003) can take many years to materialise.

- (5) Outcomes which are subjective or intangible can be difficult to capture through proxy indicators (Heinrich 2002; Smyth and Dow 1998), particularly those which are Ballung concepts (Cartwright et al. 2016). There is the danger that what matters becomes 'what's measured' (Bevan and Hood 2006), with more important outcomes not so easily reduced to measurable indicators being side-lined.

These five limitations present substantial epistemic barriers to effective outcomes management. Limitations 1-3 concern difficulties with understanding causation, while limitations 4 and 5 concern challenges in measurement and monitoring. These limitations inevitably result in a weakening of accountability, objectivity and validity which problematises rational strategic planning, prediction, monitoring and evaluation. In addition to these epistemic challenges, a set of challenges relating to control are also apparent in the literature. Thus, even if these epistemic limitations could be overcome, public managers would still face additional challenges in mounting an effective response. These include:

- (6) The creation of perverse incentives – actors cannot reasonably be held to account over factors which they cannot control, and so 'perverse incentives' are created to manipulate factors they can: skewing, distorting, or forging performance information (Lowe and Wilson 2015; van Thiel and Leeuw 2002).

- (7) The lack of control over autonomous networks – there is often no central authority with the ability to enforce new accountabilities relative to outcomes or otherwise coerce coordinated action. In public management and governance, the fragmentation of public services and the rise of networks in governance and partnership settings have increasingly necessitated working consensually through autonomous networks (Christensen and Lægreid 2007; Rhodes 1997).

The feasibility of the rationalist approach to outcomes management

The epistemic and control constraints outlined here problematise the Rationalist approach in a number of ways. It becomes difficult to mount an effective service response unless there is a clear, singular and measurable vision of outcomes (limitations 3, 4 and 5), unless their causal dynamics are well understood and a strong evidence base exists to inform rational action (limitations 1, 2 and 3), and unless there

is a central authority with the ability to enforce direct accountabilities or otherwise incentivise a coordinated response (limitations 6 and 7).

Where Rationalist approaches to outcomes management have been imposed despite these limitations, evidence shows distortions in behaviour. Firstly, the imposition of high-level outcomes can result in managers feeling as if they have been given a 'lottery ticket' (Boyne and Law 2005), and lead to resignation rather than innovation. Wimbush (2011, p.215) for instance documents a 'tendency to fall back on what data is available and what is easily measured, thereby missing the most important and relevant outcome measures'.

In situations where performance incentives are linked directly to the achievement of outcomes, a routine and extremely concerning problem is 'gaming' behaviour, as limitation 6 notes. Lowe (2013) argues actors cannot reasonably be held to account over factors which they cannot control, and so performance targets encourage the manipulation of performance data. Examples from public management scholarship are documented with striking consistency (Bevan and Hood 2006; Lowe and Wilson 2015; Smith 1995; van Thiel and Leeuw 2002), from unanticipated behavioural modifications (van Thiels and Leeuw (2002) output distortion (Bevan and Hood 2006), up to and including the outright fabrication of results (Lowe 2013; Lowe and Wilson 2015). For Lowe (2013), this is the inevitable manifestation of Goodheart's Law: any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes.

While some effectiveness has been demonstrated by 'targets and terror' (Bevan and Hamblin 2009), these instances do not account for genuine measures of outcome. Outcomes are noted by their distance from activities and even outputs, and thus present additional difficulties to measurement, planning and evaluation (Boyne and Law 2005; Hienrich 2002; Smith 1995), exacerbating perverse incentives.

Responses to the challenges discussed have fallen into two camps. The first has treated the complexity presented by outcomes as a technical challenge to be reconciled by more sophisticated (yet paradigmatically consistent) strategies, tools and models. Beefier causal models have been developed to simplify the complex effects of multi-stakeholder activities through programme theory (Weiss 1995). Pawson and Tilley's (1997) Realist Evaluation is a notable extension, developed to capture multiple pathways of causation and pay explicit attention to contextual factors. John Mayne's (2001) Contribution Analysis is another significant advancement, relaxing the focus on 'attribution' towards a rounded understanding of 'contribution' through the analysis of rival explanations for change.

Public management scholarship has largely focussed on how the 'wicked issues' (Boyne and Law 2005) of outcomes can be reconciled through better outcome indicator design and better strategic planning. For instance, Heinrich (2002) advises paying close attention to choosing indicators which are well-aligned with outcomes, which are inexpensive to administer, and which make it difficult to improve through means other than improving performance directly. Boyne and Law (2005) further address the unique measurement challenges presented by outcomes, while Mayne (2007) documents many cultural and institutional barriers to results-based management. Public management scholarship in the main has adopted this view that the inherent complexity of outcomes can be reconciled with the simplistic requirements of the Rationalist Approach.

More strident critics of the outcomes approach would contend that this reconciliatory view approach is an attempt to square the circle. Lowe (2013) for instance asserts that results-based management will *always* encourage gaming. Miller (2014) writes similarly of how the managerial 'proving' agenda behind outcomes will *always* subvert and distort their 'improving' potential. For Smyth and Dow (1998, p.291), outcomes 'promise of a semblance of order, control, and certainty', but *always* deliver the

opposite. The problems of complexity in this view are fundamentally incompatible with the simplistic worldview implicitly adopted within the Rationalist approach.

The universality of gaming (see Lowe and Wilson 2015), the poor record of key examples of the Rationalist Approach in action such as Social Impact Bonds (Disley et al. 2001), and the startling lack of positive independent accounts of Rationalist approaches such as Friedman's Outcomes-based Accountability (Lowe 2013), results-based management (Mayne 2007) or Payment-by-Results schemes, make it increasingly difficult to concur with the former camp that its problems are merely technical. While the intentions behind the Rationalist approach to outcomes management are debatable (and likely varied), it is the complexity of outcomes – their immeasurability, externality, ambiguity, and causal uncertainty – which undermines the Rationalist approach as a meaningful architecture for outcomes management in all but the most simplistic situations. Furthermore, in adopting a reconciliatory approach, we inevitably cut off a whole class of important outcomes which are too distal, too uncertain or too contested for incorporation, and therefore undercut the potential of an outcomes approach to realise transformative improvements.

The Complex Systems approach to outcomes management

The previous section has found that outcomes tend to be conceptualised and operationalised through a Rationalist approach to outcomes management which views outcomes as the results of linear service production processes. This approach was documented to encounter significant epistemic and control constraints and fall short in situations of complexity. In addition, the rationalist approach, while forming the theoretical backbone of the role of outcomes as 'results', cannot summon a theoretically robust explanation for their role as 'shared goals' or 'individual value'.

Since Smyth and Dow (1998, p.291) wrote that 'outcomes appear to have become part of a naturalised and largely uncontested discourse', it is only recently that public management scholarship has begun to engage seriously with outcomes as a

theoretical, paradigmatic and conceptual issue, rather than merely a technical problem (Lowe and Wilson 2015; Lowe et al. 2016; Lowe 2017). Partly this is because the dominance of the Rationalist approach has 'rendered other discourses irrelevant' as Smyth and Dow (1998, p.292) note, but also because alternative perspectives have not gained significant exposure in the public management literature. A viable and compelling alternative conception, developed within the social epidemiology and public health literature, has understood outcomes not as products of service production chains, but as the emergent products of complex systems.

The social determinants of outcomes

Systems interpretations of health outcomes stem back at least as far as Emile Durkheim (1897) in the late 19th century, who observed that society is more than a sum of individuals, but a specific reality with its own irreducible characteristics. Durkheim's studies of suicide rates cited in the first section of this chapter concluded that suicide was a product of social structure beyond psychological drivers, and that high rates of suicide amongst specific communities could be best explained by a relationship between individual characteristics and social norms.

More recently social epidemiologists in this tradition have argued that complex social structures create distributions of population health (Rose 2001; Kindig and Stoddart 2003). This view has also influenced policy debates and social interventions. In the UK, the Black Report showed that life expectancy and morbidities were strongly attributable to social class (Department of Health and Social Security 1980). The World Health Organisation Commission on the Social Determinants of Health has linked health inequalities to a much wider array of factors such as governance quality, social policy, social norms and wider economic trends (WHO 2008). Schensul (2009) argues in such a context that taking health outcomes seriously demands an appreciation of how such disparate factors interact across different levels to produce emergent distributions of outcomes.

Recently, population health theorists have drawn on complexity theory to articulate a view of outcomes as the *emergent products of complex systems* (Fink et al. 2016; Jayasinghe 2011; Jayasinghe 2015). Fink et al. (2016) for instance contend health 'emerge[s] from the complex interplay of health-related factors at multiple levels, from the biological to the societal level'. As an illustrative example, Figure 2 presents the findings of a 2007 exercise in mapping obesity factors was carried out by the UK Government's Foresight Programme, using a causal loop diagram to model the interconnected and nested sets of factors spanning psychological, social, economic and environmental spheres (Vandenbroeck et al. 2007).

Figure 2. The UK Government Office for Science Obesity Systems Map (Vandenbroeck et al. 2007)
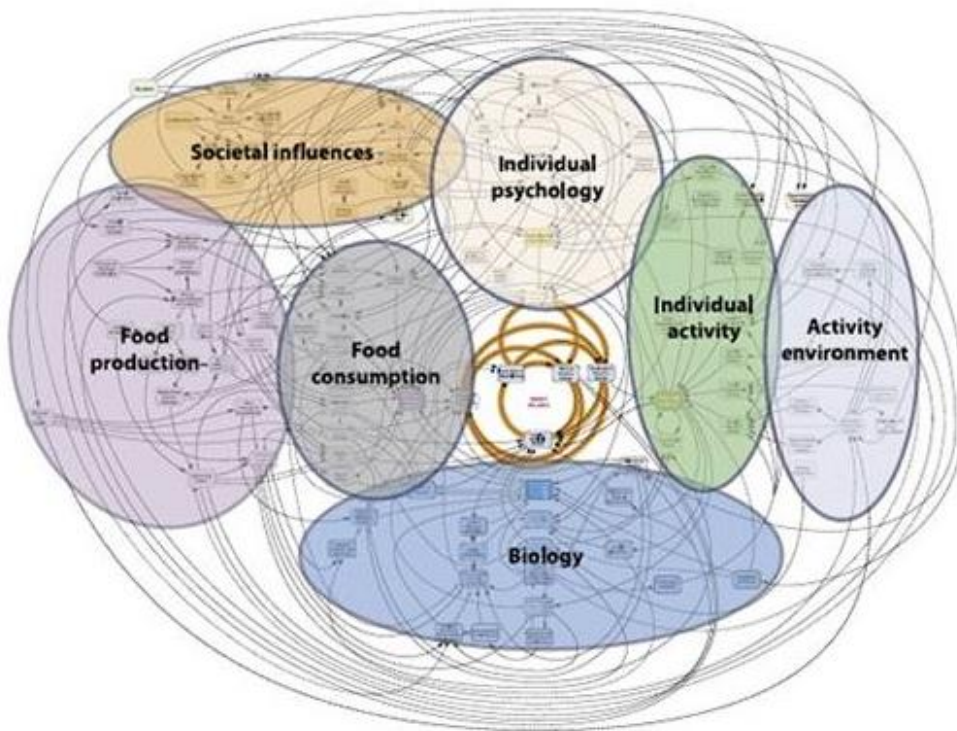
The resulting 'obesity system' map[1] visualises outcomes as emerging from a complex web of interactions between numerous nested systems. Responding to this complexity, researchers in public health, health geography and social epidemiology have increasingly drawn from the lexicon of complexity theory to reconceptualise health outcomes and their creation (Curtis and Riva 2010; Diez Roux 2011; Finegood 2011; Gatrell 2005; Jayasinghe 2015; Pearce and Merletti 2006; Fink et al. 2016).

---

[1] An interactive form of this map can be accessed at

http://www.shiftn.com/obesity/Full-Map.html

John Stuart Mill in his *A System of Logic* makes a distinction between *homopathic* effects, in which the total effect of multiple causes is reducible to the sum total of those causes, and *heteropathic* effects, such as chemical reactions, in which the sum effect is irreducible and qualitatively different to its constituent effects (Mill 1884). Emergent properties typified by Mill's heteropathy have been a central concept in systems theory (von Bertalanffy 1968), cybernetics (Wiener 1948; Ashby 1956) and Complex Adaptive Systems (CAS) theory (Holland 1992; Gell-Mann 1995). Within CAS theory, emergent properties are linked within a wider causative system, characterised by unpredictable and autonomous self-organising constituents whose sum output nevertheless takes on an inherent order (Gell-Mann 1994; Holland 1992). Prominent examples include murmurations of starlings, termite hills, patterns of urban sprawl, or the macroeconomy.

To return to the obesity case illustrated in Figure 2, if we were to remove one part of this system – e.g. some aspect of harmful individual activity – rather than just that factor being removed, the systemic response would be unpredictable to a considerable extent (Vandenbroeck et al. 2007). Similarly, new factors which appear – for instance a targeted healthy living campaign – might impact more widely across on social factors. Thus complex systems are unpredictable, dynamic, and in a constant state of flux which create unpredictable emergent patterns of system behaviour, in this case structuring outcomes on an individual and aggregate level.

The second point is that these systems are characterised by strong feedback loops arising from densely connected nature of overlapping and nested systems. Complex systems are not chaotic, being held together by self-reinforcing negative feedback loops (often termed 'attractor basins') which give the system stability and resist destabilising forces. This implies that within public policy systems, isolated interventions are unlikely to achieve substantial impact (Finegood et al. 2011). Complex systems interventions instead must aim to change the 'whole system', with the Foresight Programme report emphasising 'the need for broad and diversified

policies or strategies to change the dynamics of the system' (Vandenbroeck et al. 2007, p.8).

The re-conceptualisation of outcomes as the emergent properties of complex systems has significant implications for their organisation of public service systems. Just as Figure 2 illustrates the problem of obesity as beyond the ability of traditional health services to resolve alone, so too do educational outcomes transcend the ability of schools to achieve independently, while crime and recidivism outcomes lie beyond the jurisdiction of police and criminal justice organisations. While this view has largely been consigned to health outcomes, outcomes across the board, each placing strain on public finances (Christie Commission 2011; Mair et al. 2015), can be understood to function similarly.

Summary: theorising outcomes as emergent products of complex systems

In contrast to the Rationalist Approach which views outcomes as the 'results' or 'impacts' of public service production processes, the Complex Systems Approach views outcomes as detached from public service production processes, instead created by the complex interactions of individuals with their environments, with public agencies in a supportive role. Outcomes in this conception have a number of defining characteristics:

- They are experienced by people, and at best co-produced by public agencies.
- They are transboundary issues, unable to be improved by any public agency acting independently.
- They are causally complex, achieved through highly individual pathways and characterised by significant uncertainty.
- They have a recursive relationship with public finances, driving demand on universal services.

This locates outcomes as necessarily complex policy problems, characterised by a high degree of uncertainty, which brings them into the territory of 'wicked problems' as articulated by Rittel and Webber (1973). Like wicked problems, outcomes are amorphous goals characterised by causal complexity and uncertainty of solutions. However, they are not *entirely* wicked, failing some of Rittel and Webber's (1973) classifications. Firstly, the ability to specify outcomes implies a degree of certainty of formulation which would confound Rittel and Webber's (1973) classification. In addition, concurring with Head and Alford (2015), outcomes are likely to show sufficient stability to permit learning through trial-by-error, allowing some judgement can likely be made as to the overall and comparative effectiveness of solutions. Thus public managers and policy makers are not necessarily as lost at sea as with many 'wicked' issues such as climate change (Pollitt 2015), and indeed many approaches such as causal loop modelling (Boland and Fowler 2000), problems structuring methods (Rosenhead 2006), design thinking methodologies (Design Commission 2014), or collaborative planning (Ferlie et al. 2011) are available to grapple with the complexity that outcomes present.

Outcomes in this view also demand an extension of focus beyond cost and efficiency towards the value created for populations in common with Mark Moore's (1995) public value. However this conceptualisation of outcomes does not address the features of democratic legitimacy necessary for public value governance (Bryson et al. 2014). Instead, outcomes represent key indicators of social progress reflecting the lived experience of populations.

The characteristics outlined here distinguish outcomes from related concepts of 'impacts', 'targets', or 'value'. Outcomes instead signify a range of key social problems – for instance educational attainment, criminal recidivism, health condition incidence, poverty, social capital – which are distributed unevenly across populations. This view takes an external focus on integration with a wider system to address the root causes of outcomes, rather than introspective focus on process rationalisation, taking us back

to the starting point of this review, to John Sinclair in the context of the Enlightenment, where an outcomes-focus constitutes 'an inquiry into the state of a country, for the purpose of ascertaining the quantum of happiness enjoyed by its inhabitants' (Sinclair 1798, p. xiii).

Contrasting the Rationalist and Complex Systems approaches to outcomes management

This chapter has developed two competing conceptions of public service outcomes: the Rationalist Approach, which views outcomes as the cumulative and collective impact of public intervention, and the Complex Systems Approach which disassociates outcomes from services, viewing outcomes as the emergent product of self-organising complex systems. This section develops a better understanding of these contrasting approaches through a comparison at the ontological, epistemic and theoretical levels.

Philosophical differences

Ontology

Both conceptualisations begin from a realist ontological position that outcomes objectively exist independently of our perception. Thus although outcomes may be prioritised or achieved differently from individual to individual, outcomes are a real aggregate property at the societal level. Beyond this however, the approaches differ markedly in their ontological assumptions.

As mentioned, the Rationalist Approach assumes a Newtonian or Cartesian view of reality as determined by immutable natural laws characterised by reductionism, linearity and hierarchy. The validity of these assumptions is increasingly challenged by understandings of the behaviour of complex systems, which derives its ontological assumptions from complexity theory, and specifically CAS theory. Complexity theory for Dent (1999, p.5) constitutes an 'enhanced world view' which can generate more realistic explanations of many social and organisational phenomena which violate

these assumptions. This worldview refutes each of the ontological principles outlined by the rationalist perspective.

Holism

As emergent properties, outcomes are by definition irreducible to component parts and cannot be understood through what Rouse (2008) calls 'hierarchical deconstruction': decomposition into constituent factors which can be tackled individually through specialisation of management functions. The principle of reductionism is replaced by holism, in which causation is contextually embedded and highly uncertain. This is because outcomes are produced through the complex interaction of multiple nested systems which themselves evolve dynamically. A complex systems approach instead emphasises the entangled nature of outcomes, engulfed in feedback loops and influenced by many external factors from multiple domains. Outcomes instead come to occupy a specific reality at a particular moment of time, but changes in outcome indicators are irreducible to constituent individual contributions.

Non-linearity

Secondly, in complex systems the presence of feedback loops mean that inputs and outcomes are often disproportionate: while small inputs can have large, destabilising effects, large inputs can be dampened and result in very little change. An OECD review of Scottish school systems for instance found that, 'little of the variation in student achievement in Scotland is associated with the ways in which schools differ (…) Socio-economic status is the most important difference between individuals' (OECD 2007, p.15). As a result of both the holism and non-linearity of causation in complex systems, the link between organisational inputs and outcomes to outcomes, clear-cut and proportional in the Rationalist approach, becomes highly uncertain. Causal modelling is therefore not by itself a plausible approach within complex systems, except at the very general level (Bovaird 2014; Ryan 2007).

Heterarchy

Finally, where a rationalist perspective assumes that there is a central authority with control over the actions of other actors within the system, Instead of a central authority in control and governing by hierarchy, systems and in particular complex systems are linked by heterarchical relationships with power dispersed throughout the system. In the Rationalist perspective a central authority can direct the behaviour of lower administrative units, either through a bureaucratic function of setting rules, or through the design of incentive systems and performance management approaches.

A complex systems perspective would dispute this on two counts. Outcomes as transboundary features require a coordinated response across organisational hierarchies, each of which are likely to feature multiple and competing accountabilities. The principle of 'heterarchy' more closely characterises modern public service landscapes which have become increasingly fragmented and decentralised (Rhodes 1997) and polycentric (Ostrom 2010). The second point is that even where central authorities exist with the ability to hold actors to account through rules or targets, behaviour even in this circumstance cannot be controlled. Systems-based critiques of performance management have explained that actors routinely self-organise to subvert intended behaviours (Seddon 2003; Seddon 2008). Thus complex systems operate in environments where power is shared and actors possess the ability to self-organise independently of central authority.

Epistemology

The two approaches share an ontological realism, albeit issuing from a different understanding of how such a reality operates. Where they differ philosophically is on the epistemological level, regarding how valid and meaningful knowledge can be generated about outcomes. Outcomes are viewed in the Complex Systems approach as contested and uncertain problems (Snowden and Boone 2007), which present many of the barriers noted by Rittel and Webber (1973) in their discussion of wicked problems: they are poorly understood, ambiguous and differently interpreted, they

have highly limited evidence bases on which to devise solutions, and thus have many potential means of approaching. Even if outcomes are real in an ontological sense, from a complexity perspective we are critically limited at the epistemological level in our ability to generate universal or sufficiently detailed knowledge about them.

From a rationalist perspective, objective knowledge of the causal relationships between individual organisational processes and outcomes can be generated through causal modelling using variations of programme theory (Weiss 1995; Mayne 2001). Knowledge can be generated through positivist understandings of validity through experimental methods arranged in a 'hierarchy' with randomised controlled trials (RCTs) (or systematic reviews of RCTs) at their pinnacle, and personal experience or subjective qualitative accounts at the bottom. Objective and universal knowledge can be generated about programmes, policies or courses of action which have been 'proven' to improve outcomes. Examples include large-scale impact evaluations, or deterministic costing evaluations like Social Return on Investment which assigns monetary values to variation in outcome indicators. These approaches aim to enable policymakers, public managers and service planners to make informed, rational choices based on 'what works' in relation to target outcomes.

However, the complexity critique is not absolutist, and does not reject the validity of statistical measures of outcomes. For instance, both Rationalist and Complex Systems approaches would accept Sosu and Ellis' (2014) assertion that inequalities in attainment outcomes begin in nursery and develop through schooling, despite the inherent difficulties in measuring educational attainment. However the Complex Systems approach asserts that we are fatally constrained in *understanding* outcomes through these statistical aggregations alone: what drives them, how they interact with other factors, or what they mean for those who experience them. This deeper epistemological uncertainty is derived from two principal sources.

Firstly, the Complex Systems approach asserts that we cannot on epistemic grounds possess comprehensive knowledge about all aspects of outcomes. Instead, knowledge about outcomes is distributed across organisational boundaries and contextualised by experience and practice. In common with their functionality as shared goals, outcomes are always transboundary issues which lie across boundaries of knowledge, and thus will always create multiple valid conceptualisations based on the situated position of observers. More overt epistemological limitations, for instance the lack of a strong evidence base to inform interventions in many areas of social policy, means it will often be difficult to adjudicate with any real authority between such perspectives.

As Schensul (2009, p.242) argues, 'change toward a goal will occur faster and more effectively when synchronized and supported across levels in a social system'. Rather than pursuing any absolute truth about outcomes, framing should instead incorporate diverging perspectives to achieve a holistic understanding of complex phenomena (Fisher et al. 2016), which Jayasinghe (2011, p.2) argues must take into account 'the diversity of actors, determinants and contexts'. Where the objective of the Rationalist approach is to surmount epistemological barriers and develop a single, objective and authoritative representation of outcomes, the Complex Systems approach aims to achieve a functional consensus which entertains divergent views of an objective reality. Indeed, the value of shared outcomes in governance settings is that they enable joint activity through achieving a workable consensus in situations where it is not possible to formulate a single rational course of action.

However, the Complex Systems approach maintains that outcomes-relevant knowledge is more radically decentralised than among public service managers. 'Personal Outcomes' approaches in social care (Cook and Miller 2012) emphasise how outcomes are shaped by the 'differentiated' needs and preferences of service users (Simmons 2009; Simmons 2016), and beholden to extraneous contextual factors. Thus while outcomes are ontologically real and can be sensed in the aggregate, they are

achieved through highly individual pathways, which severely limits the efficacy of one-size-fits-all solutions, and the validity of grandstanding claims to universal knowledge.

In complex systems, patterns are repeated over time, and so we can expect views, needs, preferences and experiences to group together in emergent patterns. Therefore while we can make generalisations in complex systems, outcomes cannot be guaranteed, and thus there are intractable limits on the certainty of our knowledge. As mentioned, at the design level, Ashby's (1956) Law of Requisite Variety requires systems to at least match the complexity of user needs. It would not be possible to progress in schemes such as Self-Directed Support or Personal Budgets without harnessing, processing and implementing such knowledge into the service design, making the experiential knowledge of service users important in addressing the epistemic barriers created by outcomes implying the role of co-production (Bovaird 2007) in the generation of knowledge takes on a renewed importance.

While the Rationalist Approach privileges an 'evidence-based' approach to reform, advancing service change through the implementation of external 'evidence-based' approaches, a Complex Systems approach advocates an endogenous process of outcomes improvement, harnessing and operationalizing different perspectives dispersed across service delivery agencies and at the service user level. Thus a Complex Systems approach rejects Popper's (1969) view that reality is best understood through the scientific method in separation from direct experience, and sides instead with the instrumentalism of John Dewey (1989) in viewing experience as a crucial component of effective decision making. Epistemology is therefore fundamentally relativist and constructivist, since how outcomes are understood depends on the experience and perspective of individuals embedded within a wider social system.

The dynamic processes of causation driving outcomes within complex systems, and the multiple routes through which outcomes are achieved, while being an objective

reality, are simply too complex to be fully understood objectively. The Complex
Systems perspective combines a realist view ontology with an epistemological
constructivism, which as Chapter 4 argues in greater depth, locates the philosophical
approach firmly within the critical realist tradition (Bhaskar 2013).

Summary of key differences

The two approaches to outcomes management described in this section have
coherent ontological and epistemological positions which illuminate the implicit
assumptions of the roles played by outcomes in a public management context,
described in the first section of this chapter. These ontological, epistemological and
methodological positions are contrasted in Table 3.

Table 3. Philosophical assumptions of the rationalist and systems approach to outcomes management

	Rationalist approach	**Complex Systems approach**
Conception	Outcomes are the cumulative products of public service production processes	Outcomes are the emergent products of complex systems
Ontology	**Deterministic.** Outcomes are real, measureable, and causation is determined by a set of natural laws – hierarchy, linearity, and reductionism – which together model public intervention as a closed system	**Critical Realist.** Outcomes are determined by systems which cannot be reduced to constituent parts. Causality is highly uncertain, characterised by non-linearity, feedback loops, and instability
Epistemology	**Rationalist.** Theory and reason underpin universal assumptions about knowledge. Objective knowledge is privileged according to an evidence hierarchy epitomised by experiments and systematic reviews	**Constructivist.** Knowledge about outcomes is contested and situated within local contexts. The experiential knowledge of citizens and front-line staff are considered crucial inputs alongside traditional evidence
Theory of behaviour	**Economic rationalism.** Behaviour is governed by self-interested responses, and can be mobilised towards outcomes through the	**Bounded rationality and Heterarchy.** Systems are too complex to be knowable, and decisions must be taken in situations of considerable uncertainty. Only networks, not

	extrinsic performance incentive systems	competition or coercion, can mobilise effective service responses

The 'Rationalist' approach to outcomes management combines a determinist ontology, a rationalist epistemology, and a deterministic theory of human behaviour derived from economic rationalism. This perspective has enabled the creation of a coherent approach to outcomes management and theory building conforming to this worldview, including notably performance management frameworks such as Outcomes-based Accountability and Management by Objectives, 'league table' improvement frameworks underpinning New Labour's Modernisation programme, or recent policy innovations such as Social Impact Bonds and Payment by Results.

The Complex Systems approach is based on an entirely different worldview, combining an ontological realism with an epistemological constructivism, much in line with Bhaskar's (2013) critical realism. This worldview is combined with a realist view of human rationality and modern theories of networks and decentralisation in governance to provide a rival paradigmatic view to the Rationalist approach. However, while this approach has relevance for understanding the functionalities of outcomes as 'shared goals' and 'individualised value', and some approaches such as systems mapping (Vandenbroeck et al. 2007), it currently lacks a comparable impact on public management theory and practice. As Fink et al. (2016) note, we lack significant theoretical or empirical work understanding outcomes-focussed approaches through a systems or complexity-based perspective, or in developing practical tools to put this alternative approach into practice.

The construction of the two opposing models masks a middleground of approaches which blend elements from both. Several evaluative approaches have been developed to move beyond the reductivism and nomothetic empiricism of the evaluative

methods which Rationalist outcomes-based approaches adopt. Realist Evaluation (Pawson and Tilley 1997) advocates a systematic theory-based approach to causal analysis, rejecting universalist explanations of 'what works'. Mayne's (2001) contribution analysis recognises the limitations of attribution in situations of complexity, focussing on the elimination of rival explanations of change while stressing the importance of a strong programme theory in strengthening causal claims. Ragin's (2008) Qualitative Comparative Analysis seeks to understand causation through the systematic analysis of causal linkages among large variable sets, attempting to model emergence through identifying 'ensembles' of mechanisms.

Yet none of these approaches fit squarely within the Complex Systems paradigm. All advocate a reductivist quest for causal certainty, and understand their function as resolving the uncertainty and ambiguity brought about by complexity. The Complex Systems paradigm instead advocates local decentralised problem solving as the only appropriate response to complexity.

The two paradigmatic models are therefore polar 'ideal types', and so inevitably downplay a substantial middleground. Nevertheless, specifying the two oppositional models allows us to more explicitly articulate and understand the different assumptions underpinning service interventions. It therefore permits a more informed scholarly analysis of approaches which do blend elements of the paradigms. For example, we can understand the success of the Peterborough SIB (Ministry of Justice 2015) not as a validation of tying payment to the production of outcomes (a feature of the Rationalist Paradigm), but as creating a space for local innovation through long-term funding and focussing on experimenting with new ways of achieving a shared outcome (a feature of the Complex Systems paradigm).

Contours of a Complex Systems approach to outcomes-based management

One explanation for the dominance of the Rationalist approach within public management is that it has articulated a coherent body of theory with testable propositions, which the complex systems approach has so far lacked. The integration of perspectives from public management, public health and social epidemiology in this chapter has contributed to a viable alternative conceptualisation which is only just beginning to feed into theoretical development within public management (Lowe et al. 2016; Lowe 2017; OECD 2017).

This is perhaps most apparent within place-based approaches such as Health Action Zones, and the Total Place pilots (HM Treasury 2010), which focussed on the needs in a holistic sense of a particular place – rather than those of individual organisations. In the latter case, outcomes-based modelling has served as the basis for service redesign and system reconfiguration (Bovaird 2010; Bovaird and Loeffler 2013).

A systems view is also apparent within Scotland's post-2007 outcomes-based approach (Elvidge 2012; Housden 2014) which 'flowed from analysis of the intractability of a number of core public policy challenges' (Elvidge 2011, p.1). Scotland's reform approach recognised the need for service integration and bottom-up service reform (Christie Commission 2011), which recognised that users, rather than providers, experience outcomes (Housden 2016; Mair 2016), and that transcending NPM is necessary to realise the transformative endeavour needed to improve outcomes (Housden 2016).

The systems view can also be seen in the changing understanding of public institutions such as universities and hospitals not just as creators of educational or clinical outcomes, but as institutions with a wider civic purpose (Goddard 2009; Hambleton and Howard 2013). In the US, this has been pursued through reframing universities, local authorities and hospitals as place-based 'anchor institutions' with a broader remit as stewards of public health outcomes (Zuckerman 2013) and an array of social and economic outcomes (Dubb et al. 2013). CIPFA (2013) have recognised the

significance of this view in a public finance context, positioning outcomes as one of the three components (along with public agencies and public resource) comprising the public finance system, albeit finding that these components are often vastly misaligned.

All of these approaches have adopted the view of outcomes as detached, transboundary problems within the public sector, rejecting their integration with service processes within the Rationalist approach (Pollitt and Bouckaert 2004; Schedler and Proeller 2010). Some consistency of strategy can also be detected among recent commentary, for instance the focus on structural reconfiguration of service boundaries, and attempts to become more responsive to the differentiated needs of users and communities (Elvidge 2012; Housden 2016; OECD 2017).

There is also an emerging understanding that complexity requires a different approach to how public services organisations function, in addition to these wider structural reconfigurations. Lowe and Wilson (2016) argue that the inherent complexity of outcomes demands a new conceptual framework for the performance management of social interventions. Lowe et al. (2016, p.2) make significant theoretical contributions to articulating what this might look like, arguing that 'the purpose of a complexity-friendly [performance management regime] is to increase the adaptive capacity of the complex system under its purview'. These authors also make the important point that strategic management and performance management, necessarily sequential features in a Rationalist approach, must become integrated since 'practice must feed back into the design and architecture of the system itself' (Lowe et al. 2016, p.2).

These approaches currently lack a uniform view of human behaviour like the Rationalist approach, and at face value may appear to embrace a naïve model of purely intrinsic motivation (Bevan and Hamblin 2009; Le Grand 2003). One response to that the scope for exploiting intrinsic motivations within the public service workforce has been unduly ignored: 75% of respondents to IRISS' survey of care sector workers

(IRISS 2015) noted being motivated by a desire to make a difference, yet only 20% of those in the statutory sector reported an increase in their capacity to improve outcomes. However, a more complex incentive system blending both extrinsic and intrinsic motivators is likely required, which is an area ripe for further exploration. Lowe et al. (2016) note the potential to create competency-based horizontal accountabilities across communities of practice, which might create a 'positive error culture' (Gigerenzer 2015). The Mutual Incentives Theory (Birchall and Simmons 2004; Simmons and Birchall 2005), while developed to analyse public participation, captures the important dualism between individualistic and collectivistic incentives in a theoretically robust way, and potentially provides some means of understanding behavioural incentives within a Complex Systems approach.

It is therefore possible to sketch out some of the contours of a Complex Systems approach to outcomes management, and there is evidence that this conception is beginning to take a foothold within public service reform approaches (e.g. Housden 2016; OECD 2017). While this review demonstrates elements of this approach being applied across a range of public service areas, there remains a lack of cohesion to this loose array of perspectives and initiatives. Lowe et al. (2016) have advanced this view most significantly, arguing for an explicitly complexity-informed approach to outcomes-based performance management. Their focus on increasing the adaptive capacity of individuals to facilitate improvement is a good starting point for further theoretical development in the following chapter.

Conclusion

Despite outcomes playing a significant role in public management reforms internationally, particularly within the last 20 years, the concept of outcomes has not been subject to significant conceptual research in the public management literature. This literature review has drawn together research and conceptual development from public management, public policy, social policy, education, evaluation, social

epidemiology and public health to reveal outcomes as a contested and multi-faceted concept with key implications for the design and delivery of public services.

Drawing from public health and social epidemiology literature, two paradigmatic approaches to understanding and operationalising outcomes in a public management context are articulated: the 'Rationalist' approach, which conceptualises outcomes as the end results of linear public service production chains, and the Complex Systems approach which conceptualises outcomes as the emergent products of complex systems. The review has found that the Rationalist approach is subject to a number of key epistemic and control-related barriers which have limited its effectiveness and in practice. Despite the conceptual advantages of the Complex Systems approach, it remains theoretically underdeveloped particularly within a public management context, and useful primarily as a conceptual critique of the dominant Rationalist approach. The review has ended with a brief exploration of how the Complex Systems approach might manifest in practice, which the following chapter will explore in greater depth.

Chapter 3. Theorising a Complex Systems approach to the management of public service outcomes

Introduction

This chapter draws on Complex Adaptive Systems (CAS) theory in a constructive capacity to develop an actionable theoretical framework for outcomes-focussed service improvement. It begins with an exploration of CAS theory and its adoption within public management theory, relating this to the Complex Systems approach to outcomes management described in the previous chapter. It integrates three central components of CAS theory – self-organisation, attractor states, and distributed learning – to formulate a dynamic theoretical framework to understand and harness change and improvement in an outcomes-focussed service system. Finally, it applies this framework to the case of the Quality Improvement Collaborative, priming the empirical phase of the research. Building on the alternative complex systems-based conceptualisation of outcomes developed in Chapter 2, this chapter completes this thesis' principal theoretical contribution to Research Objective 2: *to theorise an alternative approach to outcomes management consistent with a complex systems view*.

Complex Adaptive Systems theory in a public management context

While variants of complexity theory like chaos theory and dynamical systems developed in the physical sciences, CAS theory derives from the study of biological systems (Gell-Mann 1994; Holland 1992; Holland 1995). At the most simplistic level, CAS are described by Holland (2006, p.1) as 'systems that involve many components that adapt or learn as they interact'. The autonomous behaviour of these components, called 'agents' in CAS theory, generates collective properties which are 'emergent': ordered and non-reducible to agentic behaviours. There is no unequivocal formulation of CAS, with even its founders disagreeing over an exact specification (see Gell-Mann

1994). The basic principles of the approach are clear enough however to articulate a set of cohesive assumptions governing the behaviour of agents, systems and their environments. The following list of nine propositions are synthesised from key foundational and applied texts, including early work (Gell-Mann 1994; Holland 1992; Waldrop 1992), and key applied research including Dooley (1996), Cilliers (2001) and Rouse (2008):

1) A large number of agents exist, which are densely interconnected and able to exchange information.

2) Agents are intelligent, acting according to individual 'schemata': pre-programmed 'mental maps' which guide actions based on their context. These schemata evolve through past experience, and expected future system states, but are based on subjective interpretations of reality derived from the vantage point of that individual agent. Knowledge within a system is thus dispersed and constantly evolving.

3) Agents possess substantial autonomy, and thus have the propensity to self-organise in dynamic and unpredictable ways.

4) Agents in CAS optimise performance against a shared 'performance' or 'fitness landscape', which represents their environment. This means that the agents are not entirely autonomous, rather that behaviour is determined according to environmental constraints.

CAS theory also features distinctive assumptions about behaviour at the wider system level (collectives of agents):

5) The dynamic exchange of information among agents through autonomous learning and self-organising imbues the system with a distinctive systemic memory.

6) Systems are 'open' to their environments, and system behaviour both shapes and is shaped by environmental constraints, as agents try to maximise their fitness against a changing performance landscape. This leads to the dynamic co-evolution of agents, systems and surrounding environments over time (Holland 1996).

7) Environmental perturbations can give rise to highly non-linear system behaviour. Small changes in input can spread throughout the system, amplified by positive feedback loops. Conversely, CAS can exhibit high levels of resilience, dampening the impact of significant perturbations through negative feedback.

8) System behaviour as a whole is not chaotic, but manifests in patterns and regularities repeated as fractals across the system, and develops coherent trajectories through evolution. These trajectories are determined by 'attractor' states which systems move towards over time, giving the system some degree of stability, coherence and predictability.

9) Finally, these dynamics between agents, systems and the environment produce an emergent order, which is resultant from, but not reducible to, the underlying dynamic patterns of interaction within the system.

These propositions link the behaviour at the micro (agentic) and macro (systemic) levels, with the interplay between these levels structuring overall behaviour. While agents draw upon structures and rules to inform practice, that structure itself is an emergent property of agentic interaction. In sociological terms, this approximates Giddens' (1984) theory of structuration in which agency and structure are locked in a recursive dialogue.

These characteristics of agents have significant implications for the nature of the systems that together they co-create. Most notably, the constant process of dynamic

self-organisation among agents makes the notion of equilibrium in systems an 'essentially meaningless' concept (Waldrop 1993, p.147) and implies that systems are characterised instead by 'perpetual novelty'.

CAS theory and public management

In public administration scholarship, complexity theory found early application within Douglas Kiel's early study of dissipative structures (Kiel 1989; Kiel 1993), and Comfort's (1994) study of self-organisation within public organisations. CAS theory gained a much more significant foothold during the 1990s as a description of organisations and organisational processes (Anderson et al. 1999; Dooley 1996; Stacey 1995), where it became viewed as a source of organisational effectiveness through organisational learning (Senge 1990) and adaptive leadership (Heifetz and Laurie 1997).

The fragmentation of services and increasing prominence of networks in public management from the mid-nineties onwards (Rhodes 1997; Kickert et al. 1997) created a more complex governance environment in terms of the number of actors involved, the forms it took, and the challenges it faced (Jessop 1997; Laegreid and Christensen 2013). Complexity theory has since enjoyed a significant focus in public policy (Dennard et al. 2008; Geyer and Cairney 2015; Morçöl 2013), and public management scholarship, with notable applications including within planning and strategic management (Bovaird 2008), leadership (Murphy et al. 2016), improvement and innovation (Rhodes 2013), and implementation (Butler and Allen 2008). Edited collections in leading journals (e.g. Teisman and Klijn 2008), and numerous edited books (Geyer and Cairney 2015; Rhodes et al. 2010; Teisman et al. 2009) have matured the field considerably.

Rhodes et al. (2010, p.2) contend the allure of complexity theory for public management has been to offer, 'an intellectual framework with which to observe and seek to understand, in a fresh manner, the functioning of public management systems'. However, as with the evaluation literature (Rogers 2008), complexity has

resultantly tended to be viewed as a negative force: an undesirable but unavoidable component of public management functions which subverts rational programme design and management.

In contrast, CAS theory has tended to be employed in a constructive capacity within organisational theory (Axelrod and Cohen 2000; Capra 2002; Senge 1990; Stacey 1996; Stacey 2001; Tsoukas 2005), which has long understood the links between self-organisation and innovation and creativity (Stacey 1995), and between adaptivity and the ability to respond effectively to a changing environment (Senge 1990). Similarly, its application within resource management has focussed on its utility in reaching consensus and negotiating better decisions than bureaucratic procedures (Booher and Innes 2010; Connick and Innes 2003) and in improving resilience to external shocks (Berkes et al. 2003; Olsson et al. 2004). The intention of this thesis instead is to explore CAS theory in a similarly constructive capacity in order to supersede the limitations of the Rationalist approach described in Chapter 2.

CAS theory in a constructive capacity

This purpose of this chapter is to articulate how public service systems can be informed by the dynamics of complexity to better achieve outcomes. It is useful to distinguish between CAS features which determine agentic behaviour, and those which describe the macro-level system behaviours which are resultant from these. We can then understand some of the defining features of CAS – including emergence, non-linearity, or unstable behaviour like phase shifts and bifurcations – as collective manifestations of underlying agentic processes generated through interaction with their environment. Other features discussed in the previous section, notably self-organisation, attractor states and adaptive learning, define agentic behaviours which are causative to these systemic behaviours, albeit in a way which is highly unpredictable. These latter agentic-level factors serve as the starting point for theoretical development in this chapter.

Recasting outcomes as performance attractors

In CAS, the process of self-organisation, taken in aggregate, is not entirely random, but guided by 'attractors' which underpin repeated patterns of behaviour (Gilstrap 2005). These attractors represent the system's boundaries or parameters, and structure loosely defined trajectories which systems can be observed to follow over time (Pascale et al. 2000; Wheatley 1994). Attractors have been called 'disorder organisers' which imbue the system with a semblance of stability and predictability in the long run (Dolan et al. 2003). In policy settings Haynes (2008, p.404) describes attractors as 'dominant logics and values (…) that are constantly being reinterpreted and redefined'. Bovaird (2008) meanwhile contends attractors describes as perhaps the only form of order and predictability in complex systems.

The effect of attractors is often compared to a magnetic field, actively pulling systems towards distinct trajectories (Senge 1994; Wheatley 1994; Pascale et al. 2000). Such trajectories can be deterministic, as with 'point' attractors which operate through negative feedback loops to guide the system to a 'stable position of rest (Pascale et al. 2000, 70). In contrast, 'strange attractors' exhibit dynamic behaviour, tending not towards equilibrium but encouraging chaotic patterns of bounded instability (Gilstrap 2005). Strange attractors are defined in an organisational context by (Stacey 2003, p.44) as 'patterns of behaviour, that is, shapes in space or movements over time, which are never exactly repeated but are always similar to each other'.

However, in biological systems attractors merely describe, rather than actively cause, patterns of regularity, and accordingly attractors cannot be consciously manipulated. While this holds for natural systems where the environmental parameters constraining behaviour are a function of their physical environments, within management theory and public policy, access to levers of policy, rules and regulation, and the ability to design incentive systems yields a special capacity to modify the rules of engagement. For Dolan et al. (2003, p.30), knowledge of attractors makes it possible to 'lead a system to its aimed status', a faculty which Bovaird (2008) argues gives the ability to

define the parameters within which an otherwise chaotic process of self-organisation takes place.

In organisational theory and public administration, attractors have been represented in this active sense as shared visions, which provide common goals and points of reference (Gilstrap 2005; Pedder and MacBeath 2008), which can encourage goal-directed behaviour without engaging in detailed governance (Palmberg 2009). Dolan et al. (2003) argue that 'shared values' are also powerful organisational and inter-organisational attractors. In common with 'visions', these facilitate the development and maintenance of networks and partnerships, however do so through stressing the cultural similarities between agents and creating common ground. Senge et al. (1994, p.299) argue similarly that successful learning organisations are those able to build a 'collective sense of what is important and why'.

For Gilstrap (2005, p.63), shared vision can function as a 'frame of reference where the future is unfolded within the dynamics of the organization'. A well-specified shared vision in this view is inclusive and allows everyone in the organisation to imbue a wider shared vision with an individual meaning which applies to their individual context and appeals to their values (Gilstrap 2005). This gives wriggle room for conflicting views which cumulatively can change the collective vision of the attractor (Stacey 1996). The specification of the parameters of self-organisation, allows agents the scope to innovate locally on how best to fit their environment to fit their interpretation of this vision. However in order to function effectively, any vision of the future must also link to presently-held values (Dolan et al. 2003). Stacey (1992) argues that 'vision' should be a dynamic and fluid concept which is allowed to adapt through employee agency and through discovery. Such attractors are therefore 'fluid and continually changing' (Gilstrap 2005, p.64).

Outcomes as performance attractors

The modification of attractors is one potential route whereby coordination towards shared system goals can take place. A novel approach to the management of outcomes is then to reposition outcomes as performance *attractors*, rather than performance targets. This new role shares with targets an intention to attract goal-seeking behaviour, but does so through steering and influencing behaviour, rather than holding actors directly to account against measured indicators.

Conceptually, there appears to be strong affinity between outcomes and attractors, yet such a point has not been developed in the public administration literature. The functionality of outcomes as 'shared goals' resonates strong however with the attractor role of a shared 'vision' described by Stacey (1992), Wheatley (1994), Morgan (1997), and Gilstrap (2005). The constructive role which outcomes as 'shared goals' are observed to play in governance partnerships, multi-agency social interventions, and whole-of-government approaches use outcomes within an overarching planning framework while encouraging local autonomy in their pursuit.

The attractor role of outcomes is also very apparent within the early work on personal outcomes approaches at the University of York (Ball et al. 2004; Nicolas et al. 2003; Nocon and Qureshi 1996). Ball et al. (2004, p.16) describe how over the course of a five-year project with North Lincolnshire Social Services Department, outcomes became embedded within the organisation as a 'collective sense of purpose (…) beyond the corporate froth of mission statements and glossy plans'. Outcomes then 'found substance and practical expression' (Ball et al. 2004, p.16) not as performance targets, but as the central reference point around which key organisational processes were redesigned, including performance review systems, staff training emphasis, and service processes. Consequently, Ball et al. (2004) call on public managers to make outcomes their 'big idea': almost an explicit reference to their performance attractor function.

The role of outcomes as attractors is the starting point for the development of the Complex Systems theoretical model. This gives the system a coordinative capacity appropriate for highly fragmented and decentralised public service environments, where hierarchy and direct control are impractical or impossible. This conception recasts outcomes as system 'rallying points', which can be used for two functions: to unite service systems around important service foci, and to encourage decentralised innovation and endogenous service change relevant to target outcomes.

Distributed agentic learning

Outcomes as complex phenomena are characterised by high levels of uncertainty and ambiguity (Snowden and Boone 2007). Lowe et al. (2016) argue that actors operating in such conditions must be able to make sense of changes and operate under conditions of uncertainty.

Learning in complex systems is accomplished primarily through feedback (Sterman 1994), a key process underpinning self-organisation and emergence (Richardson 2008). Feedback involves results of an action informing future actions (Eurat 2006), and is most commonly formulated in a 'cycle' of intervention, observation, reflection and action in the tradition of John Dewey (Dewey 1910), and David Kolb's (1984) theory of experiential learning. A very similar feedback process underpins the Schewart-Deming's plan-do-check-act cycle (Deming 1986) which underpins the

technical learning approaches of Quality Improvement (QI) (Shojania and Grimshaw 2005), and models of process improvement like Lean and Total Quality Management.

CAS are learning systems which are continually 'revising and rearranging their building blocks as they gain experience' (Waldrop 1993) with respect both to the behaviour of other agents and their external environment. Agents in CAS operate according to a 'schema', a cognitive structure or set of rules which determine the action an agent will take (Holland 1995; Anderson 1999). This schema forms the basis for intelligent agentic behaviour: the propensity to learn and to take predictive action, and to adjust schemata in response to new information. Schemata give CAS a systemic memory which continually informs and is informed by the learning of its component parts. This distinction between agentic and systemic learning echoes Cyert and March (1963) who argue that organisations can learn and store knowledge as an incremental improvement strategy.

The theory of organisational learning was advanced considerably by Argyris and Schön (1974), who argued individuals have a set of (often implicit) beliefs, assumptions and priorities – termed *governing variables* – which guides decision making. Learning occurs on an experiential basis when the observed results of an action do not correspond to expected results. In such a case, modifications in the actions taken can be made to correct the error which 'permits the organization to carry on its present policies or achieve its present objectives' (Argyris and Schön 1978, p.1-2). This is termed single-loop learning, which operates as the authors describe, 'like a thermostat that learns when it is too hot or too cold' (Argyris and Schön 1978, p.2). In this case, error correction is concerned with 'making techniques more efficient' (Usher and Bryant 1989, p87), and the underlying objectives, values or beliefs are operationalised rather than questioned (Argyris and Schön 1978).

With the second form in contrast, feedback affects not just the corrective response, but the 'underlying norms, policies and objectives' (Argyris and Schön 1978, p.2-3)

which constitute theories-in-use. This 'double-loop learning' modifies how problems are framed and understood, which in turn inform strategies employed. This was a qualitatively different form of learning in intent and function for Argyris (1977, p.116): 'this second and more comprehensive inquiry (...) involves questioning the role of the framing and learning systems which underlie actual goals and strategies'.

Since Argyris and Schön's (1978) contribution, the concept of organisational learning has undergone an expansive development in organisational theory (see Easterby-Smith et al. (2000) for a review). Throughout this literature, the qualitative distinction introduced between single-loop or incremental learning, and double-loop or radical learning has endured, through Watzlawick et al.'s (1974) 'first-order' and second-order change, Fiol and Lyle's (1985) lower-level and higher-level learning, and Senge's (1990) adaptive and generational learning. Argyris and Schön's (1978) distinction however has remained the most influential distinction between levels of learning, partly owing to its cogent explanation of the learning process, but also because of the novel way of linking individual and organisational learning behaviour.

Argyris and Schön (1974) explain that a private decision making framework or 'theory-in-use' guides individual action, while a publically-expressed theory of action or 'espoused theory' might suggest another course of behaviour. Organisations too exhibit this dichotomy, with learning encoded in both 'private images' and 'public maps', what (Argyris and Schön 1978, p16-17) describe as the 'media' of organisational learning, which are continually modified by individual discovery. This hinges however on information being shared and incorporated within organisational theories-in-use – most crucially from the frontline (Argyris 1977). If such a mechanism does not exist – because learning is not generated or transmitted, then 'the individual will have learned but the organization will not have done so' (Argyris and Schön 1978, p.19). This both mirrors and clarifies the nature of the dynamic relationship between agents and systems in CAS.

Learning for outcomes: the need for triple-loop learning?

Organisational learning literature extended significantly over the 1990s following Peter Senge's (1990) *The Fifth Discipline*, which positioned learning capacity as a key source of competitive advantage. Learning has also proliferated within public management literature (Gilson et al. 2009), following the Modernising Government White Paper which contended, 'The Public Service must become a learning organization' (Cabinet Office 1999, p56). Interest has been aligned primarily to the rise of benchmarking and performance management, which rely on the effective use of performance information for improvement (McAteer and Stephens 2013; Moynihan 2005). The rise of outcomes-based performance management such as Friedman's OBA and 'managing-for-results' frameworks have assumed that extending rationalist management frameworks into the realm of outcomes will increase outcomes-based learning (Lowe and Wilson 2015).

Chapter 2 has argued that many outcomes are complex problems characterised by low levels of agreement and high uncertainty (Kurtz and Snowden 2003), where causation is complex and likely only perceptible in retrospect. The organisational learning literature advocates double-loop learning where such 'deep' insight is needed: where goals are poorly understood, causal relations unknown, or the conditions of the environment change rapidly requiring revisiting such assumptions continuously (Argyris and Schön 1974; Senge 1990; Stacey 1996). This elevates the role of double-loop learning (doing the right things) over merely single-loop (doing things right), in transforming service systems towards a better alignment with desired outcomes. Double-loop learning therefore provides a potential means whereby outcomes can be pursued meaningfully through a distributed and dynamic process of outcomes-focussed learning.

In addition to single and double-loop learning, a third form, triple loop learning has been theorised which lies 'beyond and superior to' double-loop learning (Tosey et al.

2012). Where double-loop learning is taken to concern normative changes in assumptions, plans or rules, authors arguing in this tradition assert that this leaves intact the 'underlying purposes and principles' which constrain even double-loop learning2. Flood and Romm (1995) argue from a systems thinking perspective that triple-loop learning alters the authority in who gets to set aims and define rightness of purpose. Swieringa and Wierdsma (1992, p.41-42) similarly define triple-loop learning as altering the 'essential principles' underpinned overall organisational direction; while for Lassey (1998, p.11) triple-loop learning concerns changes in 'role' or 'mission', although Tosey et al. (2011) rightly criticise the position that changes in state necessarily entail changes in governing assumptions. Gilson et al. (2009), reviewing organisational learning in a public sector context, outline a form of triple-loop learning which is over and above double-loop learning, concerning questions of *why*, not *how*, public agencies pursue what they do.

The differentiation of triple and double-loop learning is problematic since neither Argyris nor Schön conceived of such a model. In some respects, the distinction between double and triple-loop learning is blurred. Argyris and Schön's (1978) original conception of double-loop learning as concerning norms, values or principles' exhibits substantial overlap with Lassey's (1998) criteria of 'mission' and Swieringa and Wierdsma 's (1992) 'essential principles' which constitute a triple-loop model. Moynihan (2005, p.204) argues in public administration context that 'the basic assumptions that underpin their mission and key policies' are altered through double-loop learning. This would push the boundaries of double-loop learning too far however for Gilson et al.'s (2009) comprehensive review in the public sector. In resource management and collaborative governance, King and Jiggins (2002) and Keen et al.

---

2 This is distinct from triple-loop learning which is conceptually aligned to Argyris and Schön's (1978) 'deutero-learning', which has been described as the capacity of learning to learn (e.g. Snell & Chak 1998).

(2005) discuss changes at the governmental level as triple-loop learning, but this is characterised by the severity of changes made, rather than a higher-order change in an underlying mental model. Surveying attempts to articulate a higher-order triple-loop learning model in this tradition, Tosey et al. (2012, p.23) conclude that the model 'remain[s] poorly defined and imprecise'.

Despite this conceptual opacity, this thesis takes the position that discussion of outcomes gives renewed emphasis to a triple-loop model. Argyris' (1977) gives the example of a private organisation cutting an entire product line as double-loop learning – analogous to service decommissioning in the public sector. However, such a decision was taken amid a set of priorities concerning shareholder value (e.g. pursuit of short-term or long-term profit) and market strategy which pre-existed and outlasted this course of action, and thus there remained intact a set of higher-level factors structuring the strategy undertaken. The very ruling of ineffectiveness in a double-loop learning process demands a preconceived and fixed notion of purpose which in public organisations is far more malleable. In a triple-loop process it is therefore the understanding and prioritisation of outcomes which is modified. Accordingly the definition of triple-loop learning adopted here is *a feedback process which modifies understanding of what desired outcomes are and how they are created.*

Self-organisation and system transformation

The ability of agents to self-organise is a defining feature of all CAS (Dooley 1996; Gell-mann 1994; Holland 1992; Stacey 1996) and has been described as their 'essential essence' (Chiva-Gomez 2003, p.105). Self-organisation refers to the faculty of agents to organise independently of central control, as they seek to optimise fitness relative to the constraints imposed by their environment. Since agents react to modifications in the strategies of other agents and changes in their environment, self-organisation in CAS is linked to a continuous process of learning and adaptation.

Two forms of self-organisation can be found within CAS. The first, called autopoietic or self-referential self-organisation (Luhmann 1995) concerns systems which reinforce or reproduce their existing structures in response to external changes (Jantsch 1980). Writing of public service and management settings, Haynes (2015) calls this 'conservative self-organisation', wherein public agencies act to entrench their policy trajectories in the face of environmental, policy or managerial attempts at initiating change. This form of self-organisation has been viewed as manifestations of self-interested behaviour aimed at maintaining power (Dunsire 1996).

The second form, 'dissipative' self-organisation, emerged through the work of the physicist Ilya Prigogine (Glansdorff and Prigogine 1971; Prigogine and Stengers 1984), who observed that systems which engaged in the rapid exchange of energy with their environments (so-called open systems) underwent rapid and irreversible change to higher states of complexity. Dissipative behaviour therefore plays a constructive role in advancing the system evolution and the production of new structures more suited to changed environmental conditions. This form of self-organisation has featured prominently in policy and public management scholarship concerning adaptive behaviour (Comfort 1994; Haynes 2015; Meerkerk et al. 2013). Haynes (2015) describes dissipative self-organisation in public administration as the creation of new rules and practices which are driven by a 'public interest' and an external focus to change a wider system.

The relevance of self-organisation for public management

Self-organising behaviour is a source of significant unpredictability as organisations do not behave according to imposed rules, laws or principles, but instead evolve endogenously through the innately unpredictable interactions of agents, and exogenously through response to changes in the environment (Teisman and Klijn 2008). Consequently, self-organisation has tended to be employed in public management scholarship as an undesirable but unavoidable phenomenon, frustrating

attempts at top-down policy implementation and managerial control. Bovaird (2008) for instance notes how self-organisation frustrated the attempts of central government imposition of Best Value, while Wheeler (2000) documents under performance management regimes that self-organisation results in data distortion rather than genuine improvement.

In the biological sciences and other social science disciplines including organisational theory and social-ecological theory, self-organisation is often seen as a key engine of innovation and competitive advantage, rather than an impediment. This has been particularly focussed on sectors of industry where there is an imperative for adaptation to remain competitive in highly volatile or rapidly evolving markets (Dooley 1996; Stacey 1996). Waldrop (1992) writes that self-organising systems display 'mutual accommodation and interdependent enterprise' which benefits both those individuals, and the wider system in which they are embedded.

Writing of self-organisation in public administration, Haynes (2015, p.41) notes it is 'best for managers to harness this creative force and to try and use it for the good of the organisation, rather than trying to suppress it'. However public management scholarship has tended to view self-organisation as a nuisance, blighting the best intentions of rational planners and strategic managers, rather than as a constructive and innovative force within public service organisations.

The relevance of self-organisation to outcomes improvement

As chapter 2 has argued, the transboundary, causally uncertain, and co-produced nature of outcomes entails that requisite knowledge for adequate comprehension is dispersed across and beyond public service organisations. Accessing and utilising this distributed knowledge is therefore essential in mounting an effective systemic response, yet epistemic and control barriers presented by outcomes make it impossible to accomplish this task through a central authority.

Self-organisation provides a means of moving forward on a decentralised basis where there is no possibility of central control or coordination. Specifically, outcomes require dissipative self-organisation, which mobilises local knowledge to produce new structures, relationships or practices. In a major study of consensus building in governance partnerships, Connick and Innes (2003, p.130) observed that CAS 'can operate effectively in a self-organizing way through distributed intelligence rather than central guidance'. Dissipative self-organisation can thus provide an engine for the transformation of services to become more relevant to the outcomes they seek to achieve.

A theoretical approach to the improvement of public service outcomes

Chapter 2 argued that dominant Rationalist Approach to outcomes management performs poorly because it attempts to reconcile the complexity of outcomes with simplistic approaches to planning and performance management. Drawing on public health, social epidemiology and recent public management scholarship, it has argued that outcomes are better understood as the emergent products of complex systems. The previous section has introduced three components - self organisation, distributed learning and attractors – which together describe how CAS adapt in concert with their environments. The following section will integrate these three components into a coherent theoretical framework which is argued to respond better to this alternative conceptualisation of outcomes.

Coordinative capacity: framing outcomes as performance attractors

Phrasing outcomes as performance attractors sets them against their more common rule as performance targets, placing focus on learning and collaboration towards a desired future system state. In place of the extrinsic incentives used to motivate outcomes focused action target based systems, outcomes perform two distinct functions. Firstly, they function as reference points for the orientation of individual action within a wider service system. Secondly, they function as rallying points,

resonating with the internally held values, principles and ambitions of individuals at different levels within service systems, prompting autonomous action aligned to the attractor state.

The technical function of performance attractors in public service improvement is therefore to encourage beneficial evolution of the system by connecting with the needs, values and ambitions of individuals within it. On a practical level strategies employed within the system about intra and inter-organisational strategy become framed in consideration of performance attractors. In systems characterised by polycentricism (Ostrom 1996), fragmentation or loose accountabilities, outcomes functioning as performance attractors can provide the only means of coordination amongst an otherwise poorly integrated service system.

Learning capacity: encouraging distributed outcomes-focussed learning

The previous section has extended Argyris and Schön's (1978) theory of organisational learning to include a triple-loop, following in the tradition of Swieringa and Wierdsma (1992), Flood and Romm (1995) and Gilson et al. (2008), concerning learning about outcomes. Figure 3 below illustrates the relationship of learning with outcomes and strategies employed to achieve them. Where the coordinative capacity afforded by attractors is intended to frame strategies and activities, learning capacity gives the system the recursive nature of CAS learning behaviour, where direction is determined dynamically through feedback with its environment.

Figure 3. Single, double and triple-loop learning in service improvement
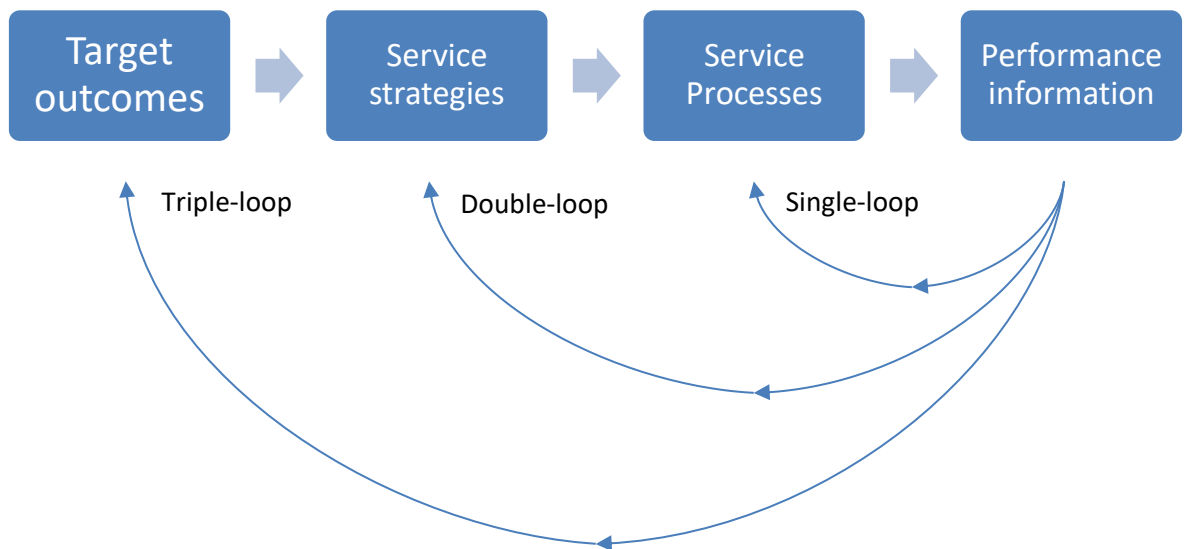
Single-loop learning concerns *doing things right* (Flood and Romm 1996). This is an output-oriented learning loop in which feedback modifies the way in which services are delivered to maximise efficiency. Reflection is based on the observation of service processes and corrective action is constrained to the modification of performative or procedural arrangements of a service, with no alteration to understanding of higher-level service objectives or strategy. An illustrative question facilitating this form of learning might be, 'how can we perform our current strategy more efficiently?'

Double-loop learning is concerned with *doing the right things* (Flood and Romm 1996). Where the focus of single-loop learning is efficiency, double-loop learning concerns a deeper understanding of impact, concerned with the alignment of outputs and

outcomes. However, this demands that the link between outputs and outcomes is sufficiently clear to be able to define an understanding of what 'effectiveness' means in a given context. In contrast to single-loop learning, the assumptions underpinning an organisational theory-in-use are modified. Impact on an organisational theory-in-use might then manifest through modifications to strategic plans, changing assumptions at a managerial or senior leadership level and alterations in espoused strategies. An illustrative question is 'could we achieve our desired service outcomes better through modifying our current strategy?'

If single-loop learning concerns *doing things right* and double-loop learning concerns *doing the right things*, then triple loop learning concerns *defining the right things*. This form of learning is focussed on the understanding of how outcomes are created and prioritised among the populations the public services serve. Learning here is detached from current service delivery and externally oriented, concerned with the myriad social forces which tie up outcomes in complex webs of causation and intractable feedback loops. In common with Elmore's (1979) Backward Mapping, this loop begins with purposive orientation as the means for strategy formulation, working backwards to consider what services should be commissioned and delivered, and what inter-organisational linkages are required.

In common with double-loop learning, triple-loop learning concerns a change in underlying theories-in-use. However, with triple-loop learning it is the premise, rather than the content, of such a theory which is challenged. At the organisational level, we might expect to see more significant organisational repurposing, the formation of new alliances, resources directed into exploring new services or ways of working. These features are summarised in Table 4.

Table 4. Descriptive summary of learning loops and their relation to service improvement

	Single	Double	Triple
Focus of learning	'doing things right'	'doing the right things'	'defining the right things'
	Linking inputs and outputs	Linking outputs and outcomes	Understanding outcomes
	Efficiency	Effectiveness	Rightness
Impact on agentic schema	Operationalised and unchanged	Modified understanding of 'what works'	Modified understanding of 'what's important'
Likely manifestations of learning	**Performative level** Modifications in service procedures Optimising delivery conditions	**Strategic level** New service creation Significant modification in service processes	**Purposive level** New inter-organisational collaborations based on shared goals Organisational repurposing typified by

		Service decomissioning	large-scale structural change
			The specification of new purposive objectives

Self-organising capacity: facilitating dissipative system transformation

Performance attractors provide agents with information about a desired course of action to rally around and orient themselves within a wider system, while distributed learning provides agents the means to make sense of their actions relative to this goal. Self-organisation then is process through which the observed disjunction between the desired and current system states are resolved, as agents seek to optimise fitness relative to the perceived constraints of their environment.

Autopoetic self-organisation describes the process of self-replication, where the current path of action is optimised and reinforced in response to external stimuli. Autopoetic systems operate through what Dunsire (1978) calls a homeostatic relationship: if an element deviates from its allocated position, it is brought into line through negative feedback and thus the system functions as a closed loop. Dissipative self-organisation in contrast transforms the system in response to external stimuli. These systems are open and highly unstable, evolving to higher levels of complexity as agents attempt to maximise fitness.

A theoretical model for outcomes improvement in complex service systems

The three components introduced here do not sit in isolation, but operate recursively to influence system change. Firstly, a system's learning capacity is moderated by its

coordinative capacity. Unless there are clear shared goals to which agents can orientate themselves and rally around collectively, there is no possibility of generating relevant performance information. The link between agentic and systemic schema change is recursive in CAS, and therefore system strategy is dynamically informed by feedback from its constituent parts, which may in turn necessitate changes in emphasis of performance attractors themselves.

Learning capacity is also intimately related to a system's capacity to self-organise. Single-loop learning from performance information is likely to promote an autopoetic process of system transformation as activities become more efficient and the execution of strategies becomes more efficient. Here, feedback gives the system predictability and stability as existing assumptions are operationalised rather than challenged. Both double and triple-loop learning prompt changes in what Argyris' (1977) term the 'governing' variables underpinning the logic of service strategies. Within these domains, learning prompts more searching dissipative changes, which in practical terms might manifest as strategic reorientation, the increasing connectedness of agents as new alliances are built, an increase in the density of information flows among agents, and significant alterations of internal service processes to better fit a new understanding of fitness. There should be instances of disruptive, not just incremental, changes in practice and over time evidence of significant transformation in the landscape of service systems.

Finally, coordinative capacity is also linked to self-organising capacity. 'Point' attractors underpin homeostatic regulation, diminishing deviant behaviour through negative feedback and so encouraging autopoesis. The specification of outcomes as performance attractors however functions as a 'strange' attractor, and attempts to initiate the dissipative transformation of systems to better suit a new performance landscape. Figure 4 below presents the final theoretical model for outcomes improvement, which integrates the function of self-organisation, learning and

attractor states to describe outcomes-focussed adaptive behaviour in complex systems.

Figure 4. The Complex Systems theoretical framework
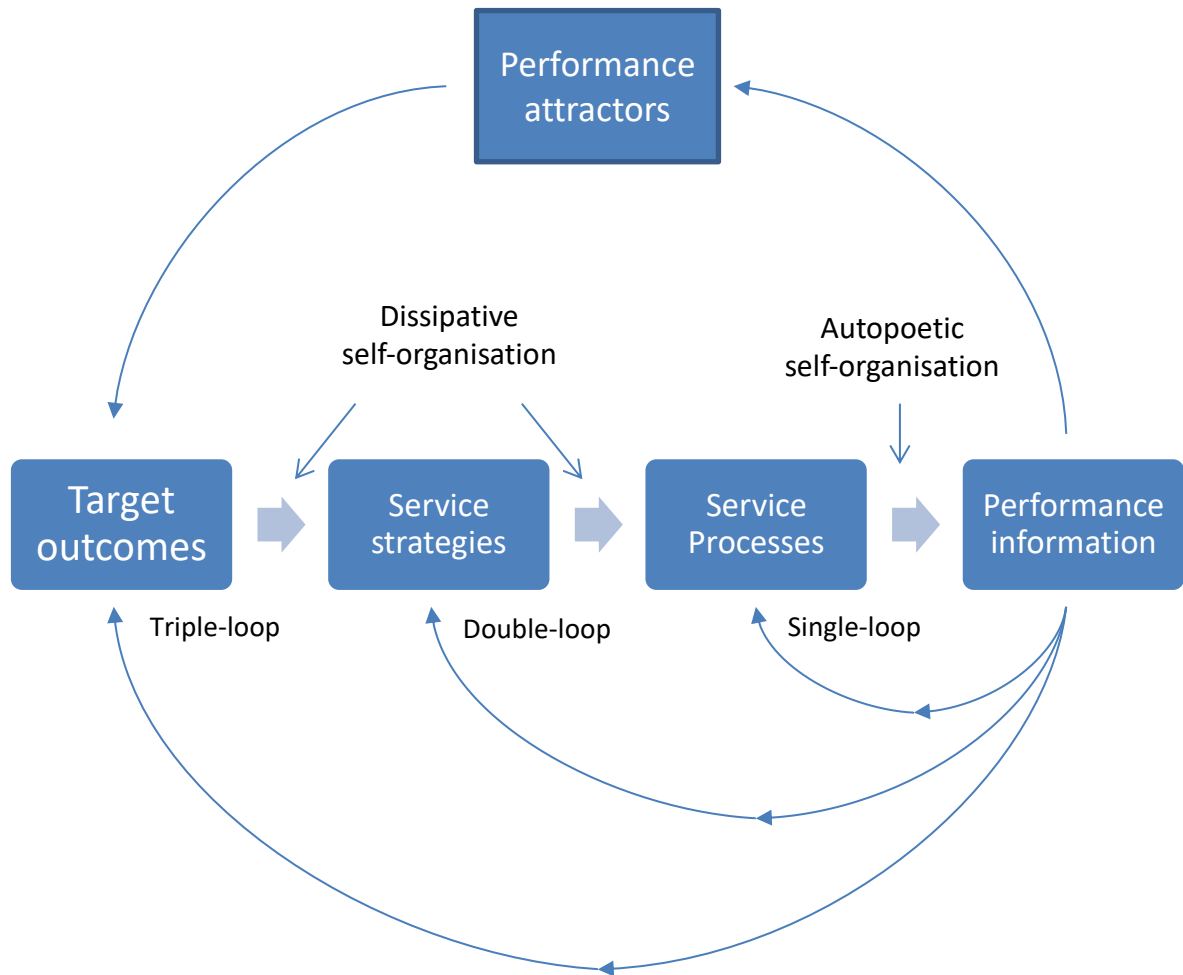
Following Lowe et al. (2016), the framework describes a process whereby a system's adaptive capacity becomes its mechanism for progress. It further delineates between three components of adaptive capacity:

- learning capacity, to explore and generate knowledge about the causative linkages between possible activities and desired outcomes

- coordinative capacity, to influence patterns of self-organisation and direct learning

- self-organising capacity, to enact a process of transformation informed by both individual learning and the system's coordinative efforts

It describes these capacities as linked through recursive feedback loops which transfer knowledge about service activity efficiency (single-loop learning), effectiveness of service strategies (double-loop learning), and the systemic formation of outcomes (triple-loop learning) from the frontline where the system interfaces with its environment and performance is generated, to other agents throughout the system.

Two distinct systemic responses can be activated by these recursive information flows. Firstly, an 'autopoetic' response is generated by the action of single-loop learning, autopoetic self-organisation, and the operation of 'point' attractors. This process results in a closed or 'homeostatic' performance improvement loop focussed on better exploitation of current strategies, and imbues the system with stability over time. Alternatively, a 'dissipative' response is generated by double and triple-loop learning, dissipative self-organisation and 'strange' attractors. This results unstable far-from-equilibrium behaviour as the system reacts to performance information to alter its structure and organisation. This latter response entails an explorative process of sensemaking where strategies and service activities become progressively aligned towards an improved understanding of target outcomes.

However, it is not necessarily the case that the dissipative response is desirable. Effectiveness (or 'fitness') within a CAS exists is not a fixed concept, but one defined through its relation to its environment. If current system performance is linked in a linear way to improvement in outcomes in its environment, the stability and predictability afforded by the autopoetic response are appropriate. As March (1991) notes, organisations always face a trade-off between knowledge-explorative behaviour and knowledge-exploitative behaviour. Where outcomes can be modelled deterministically, performance can be improved best through focussing improvement efforts on exploiting known strategies. Examples where evidence-based strategies exist include, for example, healthcare outcomes such as surgical mortality rates or the spread of hospital-related infections. Focus in such instances is best placed on improving how strategies are implemented. In this case, rational strategic planning and homeostatic performance management may be appropriate.

However, where the relation between outcomes, strategies and activities are more uncertain and poorly understood, for instance regarding criminal recidivism rates, educational attainment in secondary schools, or population health indicators like obesity, interventions are more poorly or ambiguously evidenced, and the drivers of outcomes themselves can be uncertain with multiple potential routes forward. In such cases, the relationship between service efficiency and outcomes is weak, and a knowledge-generative or dissipative response is required. As Chapter 2 has noted, genuine public service outcomes, particularly those straining public finances, are always highly complex, including persistent deprivation, health inequalities, or indicators of subjective well-being. The appropriate strategy therefore, as the Cynefin Framework (Snowden and Boone 2007) recommends, should be broached relative to the degree of uncertainty attached to its intended outcomes.

While allowing local adaptation to meet variety, the model retains a cybernetic focus on control, which can be understood as Hood et al. (2004, p.5) argues as, 'the ability of the system to keep the state of any system within some desired subset of all its

possible states'. Hood (1999) builds on the work of Andrew Dunsire (Dunsire 1978; Dunsire 1991) who pioneered the application of cybernetic concepts to regulatory functions within public administration, to outline three essential principles of any viable control system:

- A 'director', with the ability to identify a goal, target or ambition representing a desired system state
- A 'detector', with the ability to assess and report the actual system state
- An 'effector', which mandates or persuades actors to move the system state from its 'actual' position to its 'desired' position.

These three elements correspond to the three components of the theoretical framework. Performance attractors represent the role of the 'director', pursuant of a shared and mutually desired course of action. Distributed agentic learning fulfils the role of the 'detector', providing a means for local actors to sense conditions and assess suitability relative to the direction outlined. The 'effector' role in the framework is enacted through self-organisation as distributed actors take autonomous action and combine in new ways to reduce the disparity between current and desired system states.

These three elements comprise a viable cybernetic system (Beer 1984), in which the 'detector' and 'effector' functions, necessarily distinct in rational performance management systems, are fulfilled by the same organisations, and often the same individuals. This entails a form of voluntary self-regulation of behaviour, which is moderated by agents through a consideration of both systemic level goals and the particularities of their local context. However it would be a simplification to regard this as purely self-regulation, since mutuality-based approaches expose individuals to powerful 'horizontal influences' (Hood 2005, p.7). In addition, the potential exists for combinatory approaches to accountability, such as Wenger's (2010) notion of

'transversality' which combines traditional 'vertical' accountability with horizontal peer-based accountability.

Contributions to public management theory

Chapter 2 has argued outcomes invariably pose substantial coordination, control and knowledge challenges, stemming from their inherent complexity, which current approaches to outcomes-based management are not well equipped to cope with. Existing approaches have attempted to reconcile the complex nature of outcomes with Rationalist management principles as an extension of NPM logic, rather than a transcendence of it.

This framework constitutes amongst the first significant theoretical advancements in outcomes-oriented performance management, strategic management and service improvement. The principal strength of this framework is that it provides a means of tackling outcomes which responds to their innate complexity, rather than diminishes or ignores it. It provides public managers, service designers, strategic planners, and programme architects with an actionable framework to understand and manage service systems and interventions within them with a means of progressing in situations of complexity: where the factors drivers of outcomes are unclear or contested, the relationship between activities or interventions and associated outcomes poorly understood, or where the character of outcomes themselves are poorly conceived and difficult to measure.

In the Rationalist Approach to outcomes management, it is necessary to reduce outcomes to measurable indicators and to develop a causal linkage from inputs to outcomes before actors can be held to account to relevant objectives. In the Complex Systems framework, strategic and operational management functions are integrated, and strategy itself emerges from performance information (Mintzberg 1994). In this view, the pursuit of outcomes is reconceptualised as a journey through which a future system state is gradually unfolded from the present (Gilstrap 2005). Both a strength

and a challenge within this framework is that it proceeds without a detailed understanding of what such a future might look like. It relies instead on the creative energies of individuals dispersed throughout the system to co-create such a future both in vision and in practice.

In linking strategic planning and operational management, the framework aligns the with the tradition of incrementalism developed from Lindblom's (1959) work on policy design and enactment, however in drawing from complexity theory it makes explicit the potential for large-scale non-linear changes. The framework also bears similarity to Bianchi's (2016) Dynamic Performance Management, which adopts a systems dynamics approach to model the linkages between resource use and the end results of policies. That framework however remains more of a strategic management tool in common with to the causal-loop modelling used in the UK Government's Foresight programme (Vandenbroeck et al. 2007).

Perhaps the closest framework in intent is the complexity-friendly approach to outcomes-based performance management developed by Lowe and Wilson (2016) and Lowe et al. (2016), which identifies adaptive capacity as the central focus of performance improvement. Furthermore, Lowe and Wilson (2016, p.18) recognise that 'learning about practice must feed back into the design and architecture of the system itself'. Their empirical analysis concludes with a revised hypothesis, that 'improving the capacity for judgement-making in situations of uncertainty enables complex systems to adapt and meet their purpose more effectively' (Lowe et al. 2016, p14). The framework developed here shares a similar premise, its focus is instead on articulating the systemic processes which facilitate feedback and outcomes-focussed transformation.

The framework developed here is the first in public management to explicitly respond to the view of outcomes as emergent products of complex systems. In contrast to public management inquiry which have sought to reconcile the complex nature of

outcomes within existing accountability and performance management regimes (Heinrich 2002; Boyne and Law 2005; Schedler and Proeller 2010), the Complex Systems framework provides a means of moving beyond NPM in the management of outcomes. The goal here is not necessarily to supplant NPM-based approaches to public management – rather it aims to provide public managers and policymakers with the ability to meaningfully tackle the significant class of social and relational outcomes which routinely prove too complex to be reconciled within such a paradigm.

Using the framework as an evaluative tool

The framework's novelty presents a challenge to testing its utility in the planning, monitoring or evaluation of interventions, since there is no public service initiative which consciously operates based on its features and implications. Nevertheless, while there is little evidence of CAS theory being explicitly incorporated into the design of service interventions, many studies have in public management found a CAS perspective useful as a conceptual framework regardless. Applying the model as an interpretive framework in this manner makes the choice of an appropriate site crucially important for valid theoretical development and the effective testing of the framework in an evaluative capacity.

It is therefore important to clarify the criteria for its application to ensure the appropriateness of its adoption and maximise its evaluative potential. The three capacities underpinning the framework provide a clear way of specifying criteria for its adoption:

1) Systems or interventions must have articulated a clear outcomes-focus, but agents must lack formal accountability for their achievement.
2) Agents must possess significant autonomy over how they choose to organise, and be capable of forming new connections with one another.
3) Agents must be capable of generating and transmitting performance information from actions to other agents within the system.

Responding to the extra-organisational nature of outcomes, the theoretical framework is intended for operation within interventions or programmes in multi-agency settings conforming to a 'whole systems' approach. However, it is likely still of value applied within single institutions where joint goals require coordination across teams or departments which depend on voluntary engagement.

Understanding the Quality Improvement Collaborative as an outcomes-focussed intervention

As the previous section has noted, there is a paucity of public service initiatives consciously applying complexity-informed approaches consciously in their operation. Given the dominance of the Rationalist Approach in the management of outcomes, it was also challenging to find sites which meet each of the three conditions specified for the Complex Systems framework's adoption, particularly those operating at sufficient scale to tackle outcomes at the population-level. At this systemic level, there are many examples of collaborative entities, including innovation networks in governance settings (Sørensen and Torfing 2011) and forms of integration focussed on dissipative change such as Health and Social Care Partnerships in Scotland. However, few examples can be found which link planning and practice systematically in the manner outlined by the theoretical framework.

This thesis takes as its case the operation of Quality Improvement Collaboratives (QICs) – defined here as autonomous, multi-agency networks operating through a shared methodology to achieve specific shared outcomes. These are based on the Quality Improvement Collaborative (QIC) model, which combines a shared methodology based on the Quality Improvement (QI) methodology developed from the work of W. Edwards Deming (Deming 1986), and a collaborative network structure (Kilo 1998; Øvretveit et al. 2002). Following initial adoption in the USA in the late 1980s, QICs have seen increasing use in the UK healthcare system from the late-nineties within primary care and clinical healthcare settings (Ferlie and Shortell 2001).

QI was given further prominence within Lord Darzi's (2008) *Next Stage Review* which advocated QI approaches within the NHS in order to reduce waste and improve efficiency in the face of growing demand.

At the forefront of the growing international QI movement has been the Boston-based Institute for Healthcare Improvement (IHI), which has championed the application of QI methods in healthcare to achieve its IHI 'Triple Aim' of better health system outcomes, better service experiences and lower costs (Berwick and Nolan 2008). The IHI has over the past fifteen years fostered strong links with the Scottish Government and the Scottish NHS. In the latter organisation QI was systematised through the creation of NHS Quality Improvement Scotland established as a Special Health Board in 2003, which became Healthcare Improvement Scotland in 2011. In 2007, the Scottish Government and Quality Improvement Scotland partnered with the IHI to deliver the Scottish Patient Safety Programme (SPSP), a large-scale QIC aimed at reducing patient harm. Originally focused on acute hospital care, the SPSP has proceeded through two phases and now focuses on six areas: acute adult, hospital-associated infections, maternity and children, medicines, mental health and primary care.

The SPSP is based on the IHI's flagship Breakthrough Series Collaborative (BSC) model, a specific form of QIC aimed at transformational system change, which has seen widespread adoption internationally (IHI 2003; Kilo 1998). The BSC model aims for measurable improvement in defined shared outcomes within defined short (3-6 months) time periods by bringing together diverse groups of professionals to deliver collaborative improvement, and is founded on six premises (IHI 2003):

- A substantial gap between knowledge and practice exists
- There is broad variation in practice
- Examples of good practice exist, but need to be disseminated between organisations

- Collaboration between professionals working toward clear aims enables improvements
- Health care outcomes are the results of processes

The BSC model has follows a methodology called the 'Model for Improvement' (MFI) developed by IHI associates (Langley et al. 1997). The model marries the Deming-Schewart Plan-do-study act (PDSA) cycle (Deming 1986) with a set of three prior focussing questions developed (Langley et al. 1997):

1) What are we trying to accomplish?
2) How will we know that a change is an improvement?
3) What changes can we make that will result in improvement?

The four stages of the PDSA cycle represent a scientific approach of hypothesis generation and testing, providing a means to initiate, test and implement changes on a localised and decentralised basis. The PDSA cycle is often positioned as a pragmatic alternative to 'traditional' approaches to evidence-based policymaking based on a 'hierarchy of evidence' with Randomised Control Trials at the pinnacle. In contrast, PDSA cycles represent a more pragmatic approach to the generation and implementation of evidence-based, particularly in complex environments where controlling for confounding variables is impractical (Taylor et al. 2013). The stages of the PDSA cycle are summarised in Figure 5 below.

Figure 5. The Schewart-Deming PDSA cycle
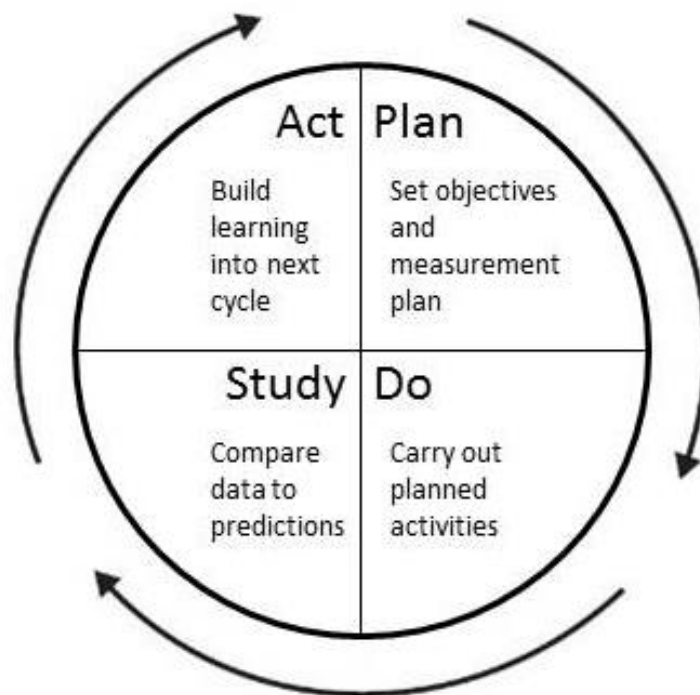
The potential for QICs to improve population outcomes, beyond achieving merely better institutional outcomes, is an area of ongoing research and policy interest (Bryk et al. 2011; Ghandour et al. 2017; Inkelas and McPherson 2015; LeMahieu et al. 2017; McPherson et al. 2015; Chief Medical Officer for Scotland 2012). Examples of outcome-focussed QICs have emerged slowly over the last decade in the US. The Magnolia Place Community Initiative, a network of 70 organisations, and the Pacoima initiative both launched in 2009 bringing together partners within defined geographical areas in highly-integrated learning collaboratives. Significant early examples in population health include the Turning Point Performance Management Excellence Collaborative in 2000, the National Public Health Performance Standards Program in 2003, and the Multi-State Learning Collaborative in 2004 (Beitsch et al. 2006). The Healthy Weight Collaborative launched in 2010 brought together health teams from 49 communities to tackled childhood obesity using the BSC model. On a multi-national scale, the IHI's 100 Million Healthier Lives project launched in 2014 also

incorporates QI tools within a loosely organised 'global innovation network' with over 1,000 members focussed around health improvement and equity of outcomes (Stiefel et al. 2016).

The success of the SPSP (thought there has been no independent evaluation) in patient safety and the close links between the IHI and the Scottish Government has led to a similar interest in applying QIC techniques at the population level (Chief Medical Officer for Scotland 2012). The launch of the Early Years Collaborative (EYC) in October 2012 expanded the BSC model to an unprecedented diversity of sectors and organisations spanning the whole of children's services within Scotland's 32 localities. The EYC became the Scottish Government's principal intervention to support the delivery of the policy objectives outlined by the high-level Early Years Taskforce (Scottish Government 2012), and one of the most prominent examples to embed the reform ambitions outlined by the Christie Commission (2011).

The Raising Attainment for All (RAFA) collaborative, mirroring the EYC's structure and methodology, was launched in June 2014, extending QI into secondary schools. The EYC and RAFA were integrated in late 2015 to form an over-arching Children and Young People's Improvement Collaborative (CYPIC) which extended QI across the lifecourse, covering children's services from pre-birth until school leaving.

The Early Years Collaborative in a Scottish policy context

The EYC was enabled by two currents in the policy agenda of Scotland. Firstly, since 2007 Scotland has pursued a whole-of-government approach to national performance management, where outcomes provide the rationale for public intervention, resource spend, and government agency priorities across all public bodies (Housden 2014). In 2011, the Christie Commission (2011) report into the future of public services in Scotland warned that 'Unless Scotland embraces a radical, new, collaborative culture throughout our public services, both budgets and provision will buckle under the strain' (Christie Commission, p.2).

The second current is the significance of quality improvement methodologies as a trajectory for service improvement within The Scottish Government and NHS Scotland. Despite no independent evaluation being conducted, the aforementioned SPSP is considered a key success story within the Scottish Government and Scottish NHS. The links between the IHI and the Scottish Government are also particularly strong, with former NHS Scotland Chief Executive Derek Feeley moving to a senior role within the IHI in 2013, and The Scottsih Government's Clinical Director, who was also pivotal in the EYC's development, becoming a senior fellow. Building on the perceived success of the SPSP, the idea to adapt QI tools to function at the population health level was first mooted in the Annual Chief Executive's Report in 2012 (Scottish Government 2012). Early minutes from the Early Years Taskforce reveal the success of the SPSP being mooted at early meetings, and feeding into the decision to adopt the model in an early years context.

The confluence of these two policy currents can be seen in the development over 2014 and 2015 of what the Scottish Government called a 'Scottish Approach' to improving and reforming public services (Cairney 2015), which aimed to put the Christie principles into practice and solidify the distinctiveness of its own policy agenda. The Scottish Approach had three components: co-production, asset-based approaches, and quality improvement methodology. The development of the EYC was considered by many as the most concrete embodiment of the Scottish Approach in practice.

Appropriateness of the case study site

As discussed, the choice of site for the empirical research is complicated by the paucity of service interventions which utilise the concepts within the theoretical framework, or draw on complexity theory more broadly. Nevertheless, the theoretical framework is primed as an evaluative model to focus empirical research on the exhibition of behaviour consistent with complexity – whether this is designed in or not. To establish the appropriate conditions for the deployment of the framework in an evaluative or

interpretive capacity, the previous section outlined 3 necessary conditions for its incorporation into empirical research. The Early Years Collaborative is a rare example of a large-scale initiative which satisfies each of these conditions, while having policy commitment to continue its work for the duration of the research project, thus providing the best available opportunity to achieve RO3. To conclude the chapter, the features of the EYC are compared with the criteria for the application of the Complex Systems theoretical framework below.

1) systems must have articulated a clear outcomes-focus, but agents must lack formal accountability to their achievement

Multi-agency QICs operate with a clear articulation of target outcomes, or 'Stretch Aims, which must be pursued voluntarily, in the absence of formal accountability, by participating agents.

2) Agents must possess significant autonomy over how they choose to organise, and be capable of forming new connections with one another.

QICs provide agents with the means of coordinated self-organisation around its Stretch Aims through the provision of local team-based forums (Workstreams) and opportunities for multi-site interaction through Learning Sessions and collaborative e-learning environments.

3) Agents must be capable of generating and transmitting performance information from actions to other agents within the system.

The MFI provides a shared methodology suitable with an orientation function (through the three framing questions), a knowledge generation function (through PDSA cycles) and a means of communication (through graphical time-series displays called 'run charts', quantitative data and storyboards).

Conclusion

This chapter has drawn on CAS theory to construct a novel theoretical framework to facilitate the development of a Complex Systems approach to outcomes management. This development constitutes the first model of service improvement which responds specifically to the conception of outcomes as emergent products of complex systems. The framework integrates the features of self-organisation, attractors and distributed learning to construct an integrated model of outcomes-focussed system transformation which enables an effective adaptive response to the innate complexity of outcomes. Three criteria are articulated to guide the selection of interventions and programmes of service reform to which the framework is applicable. The EYC is introduced and argued to be a particularly promising site for grounding the framework empirically.

Chapter 4. Research design

Introduction

This chapter sets out and justifies the thesis' research design. It begins with a description of the thesis logic and structure, introducing the three research questions tackled in Chapters 5-8. The thesis' philosophical approach is critically discussed locating its ontological and epistemological positions within a critical realist paradigm. It then describes and justifies its case study research approach and details the functioning of the research subject, the Early Years Collaborative (EYC). The chapter then moves on to a thorough explication and justification of its multiple embedded case study methodology, detailing the sampling strategy, data collection instruments, methods of analysis, and the ethical considerations made.

Research approach and logic

As Richards (2005, p.14) notes, while qualitative research is reflexive and often highly unstructured at outset, a study without a plan would be 'unacceptable for both ethical and practical reasons'. Many qualitative studies use tentative 'discovery-oriented' (Maxwell 2005, p.67) research questions (RQs) to guide inquiry, which are refined through exploration of concepts, relevant literature and theory (Agee 2009; Creswell 2007). However, as the introduction notes, this thesis faced research gaps at the conceptual and theoretical levels, which made RQs difficult to specify at outset.

Instead, the thesis in Chapter 1 set out an over-arching Research Aim to guide the logic of the thesis: to enhance the ability of public service systems and interventions to improve outcomes by advancing theoretically and empirically a complex-systems based approach to outcomes management.]. The nature of this aim lent itself more organically to the sequential achievement of four Research Objectives (ROs):

- RO1 To develop a deeper understanding of the implications which outcomes hold for public management and the design of public services interventions
- RO2 To theorise an alternative approach to outcomes management consistent with a complex systems view
- RO3 To apply this theoretical framework to an empirical case
- RO4 To draw implications from the research findings to advance public administration and social and public policy, and consider how such an agenda may be advanced in theory and practice.

The first two ROs have guided the conceptual and theoretical developments in the previous two chapters. These ROs alone however are atheoretical and detached from relevant scholarship. RQs are therefore used to locate the contributions of the research within pertinent and specific problems within public management scholarship. Three RQs are specified, which have been developed based on three considerations.

Firstly, research questions must originate from significant problems and knowledge gaps from relevant scholarship in order to contribute to issues of relevance to practice and theory, and so satisfy the thesis' requirements of relevance and originality. Each RQ articulated intends to extend inquiry within a defined area of research of importance to both specific bodies of scholarship, which are summarised in Table 5 below.

Secondly, RQs provide an over-arching structure to guide empirical analysis towards the Research Aim (Dunleavy 2003). Thus there must be a direction of travel between ambitions, questions and methods, in order for the thesis to function as a coherent whole. The Research Aim, ROs and RQs are arranged hierarchically (see Table 1) to provide a clear logic and direction of travel. Agee (2009) notes however that researchers often encounter tension between a focus on structured inquiry and signalling relevance to distinct scholarly fields. RQs 2 and 3 signal strongly to distinct

bodies of literature, however the discussion of co-production and the effectiveness of QICs in Chapters 7 and 8 integrate findings in ways which enhance the central inquiry towards the Research Aim.

Finally, RQs were developed in response to emerging empirical data. Creswell (2007, p.43) notes that often, 'questions change during the process of research to reflect an increased understanding of the problem'. Indeed, within inductive approaches such as Grounded Theory (Glaser and Strauss 2009), theory is developed entirely during fieldwork. Agee (2009) argues that RQs in qualitative research should be 'navigational tools' which also serve to 'inquire about the unexpected' (Agee 2009, p.432). This factor was particularly influential in shaping RQ 2. While the role of user feedback was a key point of interest in the development of the Complex Systems theoretical framework, it was not central to its formation. Its resultant importance during the fieldwork stage was seized upon to extend the thesis' contribution to the Research Aim while also contributing to an important research gap within co-production scholarship.

Table 5. Research questions and their contribution to scholarship

Research Question	Research gap within scholarship
RQ 1: How effective is the theoretical framework in an evaluative or interpretative capacity?	Lack of operation of Complexity Theory in a constructive capacity in public management literature (Lowe and Wilson 2016; Lowe et al. 2016).
RQ 2: How can arrangements of co-production lead to improved service outcomes, and what	Lack of theorisation over how co-production might be expected to improve outcomes

implications does this hold for a complex systems approach to outcomes management?	(Bovaird et al. 2017; Loeffler and Bovaird 2016; Voorberg et al. 2014).
RQ 3: Can the Quality Improvement Collaborative be an effective approach to outcomes-focussed improvement?	Lack of empirical studies of QICs operating in population-outcome settings, and poor appreciation of the barriers this presents (Bryk et al. 2011; Ghandour et al. 2017; Green et al. 2012; Inkelas and McPherson 2015; LeMahieu et al. 2017; McPherson et al. 2015).

Table 5 lists the three RQs and describes the research gap they address. This hierarchical structure however obscures the inherent reflexivity between aims, objectives, questions and empirical data which inevitably exist within qualitative research (Agee 2009; Bryman 2007). In reality RQs were developed in response to emerging empirical data as the scope of data constrained the ability of what the thesis could and could not claim, and as debates within literature evolved throughout the research process. RQs were refined upon completion of empirical chapters to more clearly link questions and answers, while ROs were reworded during write up to provide better focus to the thesis structure.

Finally, it is important to underscore that this thesis' contributions are resultant from both ROs and RQs. Imposed requirements for RQs to be empirically researchable mean that the thesis' preliminary conceptual and theoretical contributions to ROs 1 and 2 are not contained within discourse related to RQs; neither however are they reflected by ROs, since ROs as statements of intent lack explicit relevance to contemporary research problems. Tables 1 and 5 therefore do not reflect the thesis' full contributions, which are instead summarised in Chapter 9.

Epistemology and ontology

The philosophy of social science has been influenced significantly over the 20th and 21st centuries by two oppositional philosophical positions, positivism and constructivism. The positivist paradigm attests to an objectively knowable reality which exists independently to our perceptions of it. Objective knowledge and theory about this reality can be generated through the application of appropriate scientific research methods. These methods are commonly viewed as a 'hierarchy' with rigorously conducted randomised controlled trials and systematic reviews at the apex, and qualitative inquiry at the bottom. However, critics such as E. F. Schumacher have argued that this 'materialistic scientism' has been unable to generate useful knowledge in any other area than the natural sciences and clinical research (Schumacher 1995). The constructivist paradigm instead holds that no objective reality exists, that knowledge can only be investigated then through dialectical methods in which researcher is embedded within the social world of the researched, and that truth is bound inextricably to context. Constructivism in turn has been critiqued by positivists as obscurantist, and by realists for an absolutist perspective which limits the potential of social sciences to generate meaningful and transferable knowledge.

Pollitt (2009) criticises complexity theory for its lack of a clear ontological and epistemological stance. Despite attempts to clarify its position (see e.g. Morçöl 2013), arguments can be found positioning complexity theory as deterministic (Gell-Mann 1994) or constructivist (Cilliers 2001). This is a significant problem since a coherent philosophical position is required to devise appropriate research methods, build theory, and understand the thesis' contribution to knowledge.

The ontology of complexity

Systems theory, cybernetics, socio-technical systems and complexity theory all take as a starting point a rejection of reductionism, predictability and linearity of causation. In complex systems, the interdependent agentic relationships fundamental to causation

are dynamic and system-shaping in themselves, imparting an inherent unpredictability for overall system behaviour (Waldrop 1992). The ontology of complexity for Dent (1997) therefore constitutes an 'enhanced world view' which defies the determinism inherent within positivism. Complexity has been interpreted similarly in public administration research. Teisman (2008) for instance states actions in CAS within public administration systems are 'embedded in an immense action field full of rules, goals, expectations, routines, arrangements and histories' (Teisman 2008, p.343), making them highly responsive to context and input-sensitive. This contextual dependency for Morçöl (2013) limits the potential for generalisation and the theory-building, locating complexity within a constructivist paradigm.

However complexity theory also bears much to distinguish it from a constructivist worldview. Dent (1997) argues complexity thinking does not imply that a deterministic approach to address social phenomena is necessarily misplaced – merely that it often is. Complexity research in public policy for instance has assumed that systems can exhibit simple behaviour in many circumstances and can do so for lengthy periods of time (Geyer and Cairney; Haynes 2015). Key theorists such as Prigogine and Stengers (1984: xxiii) also understand both 'determinism and indeterminism' as valid partial explanations of reality. Thus for Dent (1997), while complexity theory supersedes the explanatory power of scientific determinism, it does not supplant it.

Complexity theory advances a general theory of behaviour linking micro-level agents to a macro-level system in which they are embedded (Holland 1992; Holland 1996; McElroy 2000). There is thus an intrinsic determinism within complexity theory which gives it generative explanatory value independent of context which holds across of a class of systems (Reed and Harvey 1992). Public management research has followed suit, on the basis that if we can understand this objective reality, we can adapt policy and management strategy around it (Bovaird 2008).

Hood (2012) notes that in a positivistic fashion, complexity theorists from the Santa Fe Institute where CAS theory developed have sought to explain social phenomena using computational methods derived from the natural sciences. Stacey (2003, p.46) has argued however that complexity within social systems is ontologically distinct since, 'human interaction is not deterministic'. Stacey (2001) argues human systems like organisations are better understood as Complex *Responsive* Processes, where the process of human relating imparts a stricter form of non-determinism than in natural or biological systems. In philosophic terms, human systems are characterised by the response of the researched, characterised by Giddens' (1984) 'double hermeneutic', also refuting the possibility of unbiased observation. Stacey's arguments have influenced other prominent complexity theorists in management, notably David Byrne (Byrne 1998), who refutes the potential of simulation-based techniques such as Agent Based Modelling, computational dynamics or big data analytics to adequately understand complex social systems.

Critical realism and complexity theory

While complexity theory does not reject the existence of an objective reality, it makes certain claims about the nature of that reality which render it distinct from both positivist and constructivist research paradigms. For Reed and Harvey (1992) and David Byrne (Byrne 1998; Byrne 2011; Byrne and Uprichard 2012), complexity theory has a natural affinity with a critical realist worldview, a perspective shared by a number of other authors (Cochran-Smith 2014; Mingers 2011; Gerrits and Verweij 2013).

Critical realism emerged through the 1970s as a critical response both to positivism and constructivism (Bhaskar 1975). Ontology within Bhaskar's critical realism is stratified across three domains: the 'real', which concerns an objective reality in which causal mechanisms which generate events, the 'actual' which concerns those events and behaviours generated by real mechanisms and which are experienced by actors,

and the 'empirical' which concerns the individuated experiences and perceptions of events. This view shares with complexity theory a focus on emergence, which echoing the Complex Systems conception of outcomes, Bhaskar explicitly relates to the generation of social problems in his later work (Bhaskar 2010).

Critical realism accords with scientific realism that theories can reflect real attributes of the world (Schwandt 1997, p.133), and thus affords the potential of theoretical validity which holds across contexts. However since researchers can only access the 'empirical' domain, and are thus reliant on perceptions, Maxwell (2008, p.164) argues 'all knowledge is partial, incomplete, and fallible', and it is not epistemologically possible to approach what Putnam (1981) has called a God's-eye view of real phenomena. While we can only access interpretations of events, critical realism is philosophically distinct from interpretivism its ontological assumptions maintain that some interpretations of reality are more valid than others (Easton 2010). Critical realists therefore employ a 'judgmental rationality' which for Danermark et al. (2002, p.10) realises the potential 'to discriminate among theories regarding their ability to inform us about external reality'.

Theories in critical realism constitute 'conjectures about mechanisms' (Ackroyd 2004, p.155), which Fleetwood and Ackroyd (2004, p.4) acknowledge can inform empirical work. However, the nature of generative 'mechanisms' in complex systems is complicated by its rejection of positivist determinism. Thus 'events' become emergent properties of complex systems which are caused by particular 'ensembles' of mechanisms (Mingers 2011), rather than the linear relationships which underpin, for instance Pawson and Tilley's (1997) Realist Evaluation. Reed and Harvey (1992) and Harvey and Reed (1996) integrate critical realist view of 'stratified' ontology with complexity theory to form a 'complex realism'. David Byrne (Byrne 1998; Byrne 2011a; Byrne and Callaghan 2013) has substantially extended this view into an actionable framework, which integrates Rihoux and Ragin's (2009) Qualitative Comparative

Analysis to systematically assess the causative potential of these ensemble mechanisms.

Integrating critical realism and complexity theory moves towards a cohesive approach to theory building. Generalising knowledge across contexts and building theory is ontologically permissible however conforms to a critical rather than a scientific realism: theories are always representations and abstractions of an underlying reality, and are therefore provisional and open to amendment or falsification. Causality will be 'neither governed by general rules nor by pure idiosyncrasies' (Teisman and Gerrits 2014, p.21), and must be understood through the interaction of factors, which may not hold over time.

In answer to Pollitt's (2009) call for philosophical clarity when adopting complexity theory, this thesis' position is that complexity theory follows critical realism in combining an ontological realism with an epistemological constructivism: even though there exists an objective reality (whose nature is augmented by complex dynamics), knowledge about reality is always incomplete, imperfect and subjective. While some have attempted to position complexity theory as an overarching 'meta-theory' in its own right (Morçöl 2013), the parallels with critical realism are so strong for Gerrits and Verwiej (2013, p.167) they argue, 'any effort to research social complexity is implicitly or explicitly informed by [critical realism]'. Locating the study within a critical realist paradigm positions it within a recognised research tradition with clear implications for theory development and empirical research.

Case Study Research

Varieties of complexity in the social sciences

While complexity theory has been influential within the social sciences, a number of authors (and indeed its advocates) have been critical of the application of a science developed within the natural and physical sciences to social phenomena (Byrne and

Callaghan 2014; Cilliers 2001; Stacey 2001). In public management scholarship, applications of complexity theory have tended to lack consistency and rigour (Klijn 2008), and have failed to articulate a consistent ontological or epistemological basis (Pollitt 2009).

Some schools of complexity, particularly the Santa Fe Institute, have advocated the direct transfer of insights complex adaptive biological systems to human social systems. In this view, a network of homogenous self-organising agents obeying 'simple rules' produce emergent phenomena which are unpredictable in their detail, but which can still be modelled or simulated. Methods such as agent-based modelling, non-linear statistical analysis, and static systems dynamics models like 'systems maps', all of which have gained prominence as complexity-consistent methods of evaluation within policy systems. While these eschewed the deterministic outlook of classical 'Newtonian' science, they still retain an innate determinism which enables prediction, statistical modelling, and visualisation of dynamic and attaches complexity as Morin (2006, p.6) describes, 'as a kind of wagon behind the truth locomotive'.

Other interpretations of complexity have noted the distinctiveness of human systems, and cautioned against uncritical transfer of the theories, concepts and language of complexity (Cilliers 2001; Stacey 2001). In this view, human systems are distinctive since agency is system-altering in a way which biological organisms and natural phenomena are not, and the structures of governance and social relations confound the view of self-organisation as autonomous and organic. This alternative sense of complexity is a 'generalised' complexity (Morin 2006), in which a 'strong' emergence is not reducible merely to the autonomous interactions of constituent agents, and therefore cannot be straightforwardly simulated.

Complexity in this thesis is informed by the view of strong (irreducible) emergence and generalised complexity (Morin 2006; Byrne and Callaghan 2014), and notes the significance of self-organisation in spite of hierarchy, rather than in the absence of it

(Cairney and Geyer 2015). Nevertheless, it maintains that the concepts and language of complexity theory are still relevant in this new context as metaphor and perspective. To return to the Obesity Map, following an intervention such as a sugar tax, we would expect the system to reconfigure itself in a largely unpredictable way, rather than result in a proportional decline in the target outcome.

Traditions in researching complex human systems

As Teisman and Gerrits (2014, p.21) note 'not every method is well-suited for dealing with the synchronic emergence that drives social reality'. Perhaps accordingly, there is still a paucity of empirical research in public administration applying complexity theory in a consistent manner (Pollitt 2009).

Associated with Complexity theory are a rich array of potential methodologies from computational modelling, statistical methods and qualitative inquiry (e.g. Morçöl 2013), some of which have made the transition to the study of human systems (Byrne 2011a). These however ignore the interrelationship of variables within complex systems which evolve over time (Blackman 2013; Stroup 1997), and rely on a problematic equivalence between the behaviour of human 'agents', and those of simple biological organisms. Complexity-based approaches such as Developmental Evaluation (Patton 1994) or Action Research (Carr and Kemmis 1983) hold comparative advantages relative these approaches in situations of high uncertainty, however are particularly intensive, and require extensive planning to facilitate access and integrate with operational management.

Other approaches such as Mayne's (2001) Contribution Analysis and Pawson and Tilley's (1997) Realist Evaluation and allow for the absorption of complexity through alternative causal pathways, however both similarly require that a developed theoretical understanding before analysis and focus on reductionism which sits uneasily with the epistemology of complexity theory.

Complexity is understood to challenge the use of programme theory (Rogers 2009) and theories-of-change approaches as Barnes et al. (2004, p.13) note: 'our experience of evaluating [Health Action Zones] leads us to suggest that this evaluation stretches the application of "Theories of Change" to a point at which it becomes both methodologically and theoretically fragile.'

Of note in this regard is Pawson and Tilley's (1997) Realist Evaluation which has gained prominence over the last ten years as a complexity-consistent evaluative framework (Marchal et al., 2012). Realist Evaluation rejects the universalist explanations of causation which the Rationalist Approach seeks to generate by developing clear hypotheses about where, how and for whom interventions work. It considers social interventions and other service 'programmes' as theories incarnate, and so considers the deployment of such interventions cognate to hypothesis testing. By acknowledging the interconnection between interventions and their environment, some have argued that a realist approach is a complexity-consistent evaluative approach (e.g. Marchal et al., 2010).

Pawson himself considered it a superior method to systems thinking, which he argued 'seem to embrace solipsism and deny that we can learn from inquiry to inquiry' and thus 'multiplies rather than solves the complexity burden' (Pawson and Tilley 1997, p.55). Instead, the evaluative approach of Realist Evaluation is to engage in meaningful reductionism (Jessop 1997); identifying key context-mechanism-outcome configurations in which target outcomes arise through particular (intended or unintended) programme mechanisms which interact with particular contexts (*inter alia*, temporal, spatial, cultural and social).

Complexity theory applied to social systems is argued in this thesis to be congruent with a critical realist philosophy, as articulated by Reed and Harvey (1992) and later Byrne and Callaghan (2014).Realist Evaluation shares this philosophical root, and also features a number of conceptual similarities. For example, Realist Evaluation's focus

on contextual interaction mirrors the importance of environmental conditions in complexity theory. Both are also linked by a rejection of universalism and simple reductionism which is argued to underpin the Rationalist Paradigm of outcomes-based public management.

Others including Sanderson (2002, p. 10) are more critical, arguing that Realist Evaluation's focus on mechanisms and contexts remains reductionist at its core. At the ontological level, Byrne and Callaghan (2014) note complexity requires not only a 'stratified' ontology but one which captures the nature of complex behaviour. While realism can be philosophically aligned with complexity theory, it lacks the language or concepts to adequately describe change in complex systems.

Conceptually, a complexity lens would also blur the sharp divide maintained in Realist Evaluation between 'context' and 'mechanism', and what lies within or outwith a service 'programme'. The conceptualisation of outcomes as emergent products of complex systems makes no such distinction between context and mechanism. Instead, configurations of factors running between and across nested systems determine outcomes.

Finally, Realist Evaluation and complexity theory have different implications for evaluative practice. In seeking out succinct deterministic configurations of contexts and mechanisms, Realist Evaluation advocates a reductionist evaluative inquiry which seeks to unearth universal (though locally-limited) truths. In complexity theory, this is not a productive avenue of inquiry, since the scale of complexity attached to outcomes is likely too great for universalist explanations of change to have any useful relevance. The alternative approach, of empowering local actors to interpret and respond to local complexity while abandoning attempts to generate central understandings of causation – is a more effective response ins situations of complexity. In advocating an essentially reductivist search for configurations of mechanisms and context, Realist Evaluation is functionally (though not philosophically) opposed to the Complex

Systems Approach, which instead deals with uncertainty through enhancing the adaptive capacity of local actors.

Validity and reliability in case study research

The case study as a structured inquiry into a bounded phenomenon is a foundational method for undertaking theoretically-informed social scientific research (Flyvberg 2006; Stake 1995; Yin 2009). Methodologists agree that case study method has a particular strength in holistic, in-depth investigation (Merriam 1998; Stake 1995; Yin 2009). By investigating phenomena embedded within a naturalistic context, case study research can be particularly effective in accessing the 'black box' of complex causal relations (Bromley 1990; Yin 2009), and thus has a history within public administration research adopting systems and complexity theories (Anaf et al. 2007; Anderson et al. 2005; Rhodes 2003; Rhodes et al. 2010) and within critical realism (Easton 2010).

Case study research has been criticised by positivists for lacking both internal and external validity. As an observation of one or a small number of cases, case study research cannot make inferences to wider populations through statistical generalisation. The reliance on perception data and the lack of a control group are also threats to internal validity since they open up a range of potential biases, including prominently confirmation bias, perception bias, and selection bias (Flyvberg 2006). For positivist methodologists such as Yin (2009), well-specified case studies can mitigate these issues through a replication of the scientific method. Internal validity can be attained through rigorous specification of case studies in their sampling criteria, research design and through addressing rival hypotheses within research (Yin 2009). External validity can be attained through replicating findings over a number of similar cases – each constituting a separate experiment, rather than a statistical observation (Yin 2009).

For others, including Stake (1995) and Merriam (1998), it is the premise of the positivist critique, rather than its logic, which is rejected. Case studies for these theorists are rooted in a constructivist epistemology which maintains knowledge is situated and inseparable from context, and therefore a social construction (Merriam 1998). The case study approach therefore rejects positivistic conceptions of validity and reliability (Golafshani 2003; Lincoln and Guba 1985) and many of its principles including context-free generalisation. Instead, the concepts of reliability and validity are supplanted by context-specific understandings of credibility and trustworthiness (Lincoln and Guba 1985). Morse et al. (2002) write that this shifts responsibility for study quality onto the researcher and the research process, rather than merely the research design.

Critical realism and case study research

Case studies are adopted in the critical realist tradition to understand the meaning which actors attach to events (Ackroyd and Fleetwood 2000; Easton 2010), ceding to the constructivist perspective that objectivity 'in the study of human affairs is impossible' (Fleetwood and Ackroyd 2004, p.3). Byrne (2011a, p.134) criticises Yin's (2009) positivistic standpoint arguing it '[remains] trapped in a variable centred understanding of causation'. Critical realism instead demands for Maxwell, (2015, p.91) a 'process-oriented view of causality' which emphasises the 'importance of context and particular understanding, rather than focusing entirely on general conclusions and laws'.

Case studies allow the exploration of a phenomenon from multiple perspectives, allowing 'multiple facets of the phenomenon to be revealed and understood' (Baxter and Jack 2008, p.544). The fallibility of knowledge requires seeking alternative explanations from multiple perspectives a key measure of study reliability (Stake 1995, p.107), and as with the constructivist paradigm, the researcher must as faithfully as possible people's understandings of phenomena, becoming 'interpreters and

gatherers of interpretations' (Yazan 2015, p.137). Thus for Easton (2010), case studies must be designed to capture the nature of change in as much depth as possible to achieve a multi-faceted and highly contextualised understanding, which has greater bearing upon reliability than assessing regularities across many different cases.

Teisman and Gerrits (2014, p.18) argue case studies of complex systems must be able to deal with, 'large amounts of data that are not coherent and countable in a direct sense and with relations that reach beyond simple takes on causality'. For Cochran-Smith et al. (2014), integrating complexity theory within the critical realist paradigm allows better contextual understanding of intersecting and nested systems, and can thus contribute to an improved explanatory account of complex systems change.

However, as Haynes (2015) notes, complexity can take time to manifest and so case studies should incorporate lengthy temporal dimensions to allow complexity dynamics to take effect and manifest in emergent patterns. Haynes (2015) also argues for a focus on the effects of a system's environment, to capture the interaction of case study systems with their external contexts (Haynes 2015). Thus a case study analysis of complex systems within the critical realism paradigm entails a resource-intensive data collection process across large timescales and multiple system levels. The case study definition adopted here therefore follows Easton's (2010, p.199) definition: 'a small number of social entities or situations about which data are collected using multiple sources of data and developing a holistic description through an iterative research process'.

Case study and theory development

Causal explanation in critical realism is generally sought through the isolation of real 'mechanisms' which govern the creation of events and comprise the highest level of causal knowledge about reality. As Sayer (2004, p.9) notes, all theorists by nature acknowledge the independence of constructs and knowledge. Theories are

'conjectures about mechanisms' (Ackroyd 2004, p.155), which Fleetwood and Ackroyd (2004, p.4) acknowledge can 'inform effective empirical work'.

Complexity theory is noted by Easton (2010) as structuring inquiry into the identification of real 'generative' mechanisms. Byrne (2011a) characterises generative mechanisms as the drivers of complex systems, whose expression represents the 'actual' ontological domain. Thus the processes of complexity theory are understood to structure the relationship between the 'real' and the 'actual' ontological domains.

The theoretical model postulated in Chapter 3 takes a process-oriented view of change and adaptation, and is not therefore suited to Byrne's (2011a) case-based approach to the determination of causation. Instead the framework speculates on potential causal relationships which determine adaptive capacity which can be argued in this instance to represent generative mechanisms. Thus we hypothesise that outcomes functioning as performance attractors will influence patterns of self-organisation among constituent agents; that distributed learning will influence decisions about strategy and orientation as new more effective strategies are learned and communicated; and that the sum contribution of learning can influence systemic understandings of outcome systems and augment the specification and presentation of performance attractors.

The framework also posits that two loops around this model are possible. Firstly, an autopoetic process of system transformation, characterised by autopoetic self-organisation, point attractors, and single-loop learning. This loop is tied to stability assured by negative feedback loops, and represented by improved service efficiency, and an entrenchment of service trajectories. Secondly, a dissipative process of transformation, characterised by dissipative self-organisation, 'strange' attractors, and double and triple-loop learning. This second loop is tied to non-linear and path-breaking behaviour, manifested through increasing connectivity, significant alterations in service strategy, and divergences over time in system trajectory.

However, in line with critical realist tradition these mechanisms and their anticipated events are provisional and 'corrigible through observation' (Ackroyd 2004, p.155). Critical realists typically employ a form of 'retroductive' reasoning (Sayer 1992) to causal analysis, based on Charles Peirce's principle of abduction (Peirce 1957). This approach seeks parsimony of explanation through iterating back and forth between theory and data (Sayer 2004). Thus theoretically-informed research must avoid the tunnel vision of a purely deductive approach to empirical research through being open to emergent outcomes and alternative causal explanations.

RQ1 is directly focussed on addressing this theoretical nexus, not to prove theoretical assumptions, but to improve the model through empirical investigation. This will be accomplished by contextualising the theoretical linkages described as theory interacts with the real world. A satisfactory response to this RQ therefore is not in the affirmative or negative, but in how the framework can be adapted and improved through contextually-situated analysis.

Specification of the case study approach

Yin's (2009) focus on rigour and transparency of process is shared by constructivist theorists like Guba (1981) and methodologists like Stake (1995). Choices taken regarding the specification of the case study must then be clear and justifiable, regardless of the philosophical position of the researcher. Considerations of transparency and trustworthiness therefore guides the disquisition of the case study methodology employed, which comprises the remainder of this chapter and is presented in three parts:

- The specification of case study boundaries.
- The description of the case study approach and sampling strategy employed.
- The research methods employed, including ethical considerations and the data analysis process.

Defining the case study parameters

EYC structure at the CPP level

In Scotland, most children's services are funded and delivered at the locality level by statutory agencies across unitary authorities, led by 32 Community Planning Partnerships (CPPs) and 13 Health Boards. CPPs were therefore recognised by the Taskforce in its 2012 Vision Statement (Scottish Government 2012) as the principal delivery vehicle for achieving the ambitions of the Early Years Framework. The Children and Young People (Scotland) Act 2014 introduced a new responsibility for CPPs to report annually on progress across these domains while CPPs are also tasked with preparing and delivering against Integrated Children's Service Plans for the locality up until at least 2018. CPPs are also scrutinised through the Care Inspectorate's external joint inspections of children's services until 2017. Inspections are intensive and command significant within CPPs, with the requirement that CPPs publish and carry through an action plan in response to the summative inspection report.

Despite these external accountabilities, CPPs maintain considerable autonomy over the structure and focus of children's services. The 2007 Concordat between central and local government established a new consensual politics based around shared outcomes specified in Single Outcome Agreements (Scottish Government 2007). CPPs could choose largely how agreed outcomes would be tackled, under the recognition that the unique local contexts which CPPs faced required local solutions. Children's services within CPPs are led by a sub-partnership body, often called an 'integrated children's services' (ICS) partnership.

ICS partnerships vary in scope and structure across CPPs, but all comprise a high-level multi-agency partnership, often a formal board, charged with delivery against children's services outcomes, and accountable to the CPP and council chief executive. Infrastructure beneath ICS boards can vary drastically, responding to both the

disparate geographical make-up of Scottish CPPs, and thematic and issue-based priorities, which often make up idiosyncratic thematic working groups accountable to ICS boards. ICS plans published by CPPs are therefore dually accountable to the CPP and its senior leadership, and to the Scottish Government.

Thus while the EYC was coordinated and supported by central government, it was recognised that leadership and operational management had to come from the CPP level. In developing the EYC, the Scottish Government approached council chief executives to sign up to its delivery on a voluntary basis. CPPs were asked to recruit an Away Team and a Home Team to fit with the BSC's model of Action Periods and Learning Sessions. Away Teams were responsible for attending Learning Sessions and bringing back insight to the Home Team. CPPs recruited the Away Team from existing ICS partnership infrastructure, however were expected to build Home Team membership progressively to extend the EYC's reach into the wider early years workforce. Members of the Home Team were responsible for initiating tests of change (TOCs), and adopting or transferring learning outwith their individual service settings. CPP EYC teams would also organise Home Team events, full or half-day functions to bring the whole CPP team together, share learning and deliver training.

These TOCs aim to develop evidence of improvement through making small changes to service practices and documenting their effects through adoption of the Model for Improvement (MFI) described in more depth in Chapter 3. Improvements can then be taken to scale at low risk, scaling to larger case loads, from one practitioner to several, across service teams and locations, or moving to other sites or organisations entirely. While there is also scope for incorporating qualitative data, the MFI is perceived as a primarily quantitative tool predicated on a measurable dependent variable. Changes can be determined to be an improvement through using 'run charts', where an independent variable is plotted against a dependent variable before and after an intervention. A service change can be determined to be an improvement if documented through consecutive observations following an intervention, theoretically

allowing the separation of natural variation from intervention effects. TOC leaders can come from all levels of organisations, senior staff to frontline practitioners, and can be led by a single individual or a wider team.

CPPs were also asked to set up four multi-agency 'Workstreams', three of which were expected to lead collaborative testing and coordinate work towards each Stretch Aim, while the fourth, a 'leadership' Workstream, was intended to provide strategic leadership for the EYC more generally. Workstreams were appointed a 'lead' and a 'depute lead' from senior leadership within relevant children's services to coordinate activity within the Workstream and manage Workstream meetings.

Finally, CPPs were asked to staff EYCs with a named Champion, a Programme Manager and a Data Manager. The EYC Champion was intended to be a senior figure within the CPP – a senior children's service manager, prominent elected member, or CPP leader – to act as the local figurehead, and provide visibility and strategic impetus within the CPP. While CPPs were initially expected to coordinate the operational management of the EYC without a dedicated leader, CPPs were asked as soon as possible to appoint a Programme Manager (PM) take over. Most CPPs appear to have appointed dedicated PMs at around Learning Session 3. Finally, a data manager would take responsibility for ensuring that real-time data is available to support improvement projects, and to assist TOC leaders with the measurement and management of data, however only some CPPs appear to have appointed or kept on data managers.

Of these roles the PM was most prominent and important. The PM tended to take responsibility for the coordination and facilitation of the EYC structure and operations on a day-to-day basis, with responsibility for both strategic development and technical project assistance. PMs functioned as the link between strategic and operational management, and the Home and Away Teams. They also became the main point of contact for the Scottish Government's Practice Development Team under the Early Years Quality Improvement Unit (EYQIU) and the Scottish Government Improvement

Advisors who would later replace them, although they remained formally accountable to the CPP rather than the Scottish Government. PMs were also in communication with one another, through semi-formal PM networks, through the Extranet, and through informal regional PM forums.

Framing the EYC as a case study

This section has described the complicated structure of the EYC as stretching roles and accountabilities across multiple levels of the children's services system. To aid the reader in understanding the EYC's operations, it is helpful to distinguish between two functionalities: a 'strategic level' where key managerial decisions are made concerning strategic direction at the National or CPP level, and an 'operational level' in which testing and developing TOCs, and sharing and adopting changes elsewhere are conducted. These two functionalities provide two interlinking units of analysis, which together permit the construction of responses to all RQs. Figure 6 provides a diagrammatic explanation of the EYC's infrastructure, while Table 6 outlines the key roles, responsibilities and accountabilities of partners involved in taking it forward.

Figure 6. Structure and accountabilities in the Early Years Collaborative

CPP strategic landscape

Integrated Children's

Community Planning

Scottish Government

Early Years Taskforce

Early Years Unit

Practice Development

CPP-level EYC structure

Programme

Workstreams

Away Team

Home Team

External Children's Service Organisations

Strategic-level — Formal accountability

Operational-level — — ► Informal accountability
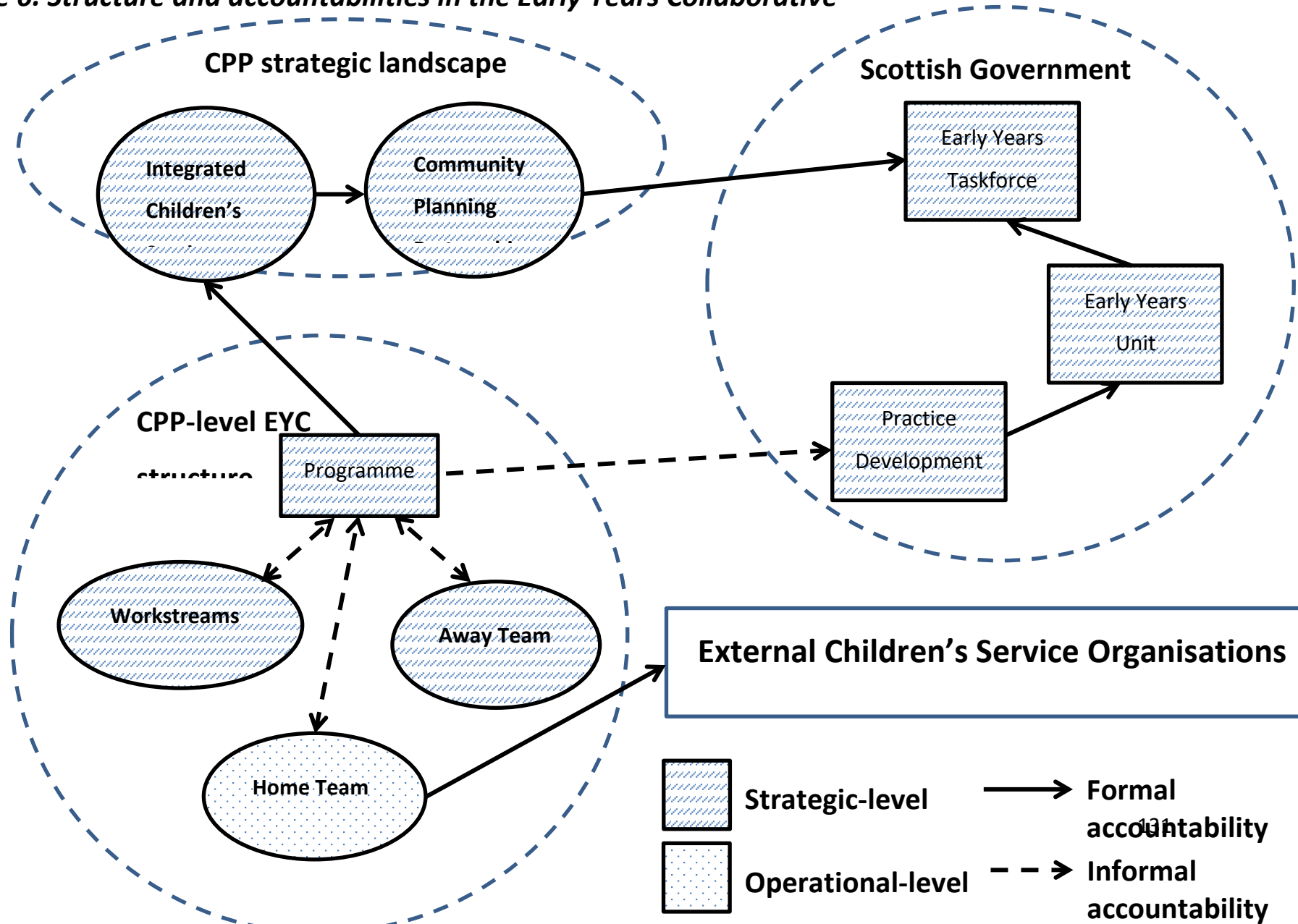
Table 6. Roles, responsibilities and accountabilities across the EYC

	Title	Role within EYC	Accountability
Strategic-national level	Early Years Quality Improvement Unit	Responsible for development, strategic and operational management of the EYC	Early Years Taskforce / Children and Families Directorate, Scottish Government
	Early Years Taskforce	Early Years Taskforce was established in November 2011 to oversee the implementation of the Early Years Framework and related policy, including the EYC	The Taskforce has overall responsibility for the Early Years Change Fund and the development of the Early Years Collaborative.
	Quality Unit	Responsible for technical support and jointly developing the EYC with the Early Years Quality Improvement Unit .The Clinical Director took a leadership role throughout the EYC's lifecourse	Health Directorate, Scottish Government
	Practice Development Team	Set up to facilitate conversion of Early Years Framework principles to action. The Practice Development Team advise CPPs and support the delivery of the Change Fund and took on responsibility for technical support for delivery of the EYC	Early Years Taskforce / Children and Families Directorate, Scottish Government
Strategic-CPP level	Champion	Responsible for 'championing' EYC at a senior level within the CPP, ensuring that it remains a strategic priority at the highest level, and providing visible leadership for the wider early years workforce	CPP and Chief Executive
	Programme Manager	Overall strategic management responsibilities: - supporting TOCs with technical and project support	ICS and its sub-partnerships, with expectations to report to

		- communication and networking responsibilities to build and sustain EYC membership - representing the EYC within ICS and reporting on progress	and engage with the Scottish Government
	Workstream leads / deputes	Setting agenda and coordinating discussion in Workstream meetings, providing local and sectoral leadership, ensuring the innovatory and collaborative capacity of Workstreams	Informally accountable to ICS and relevant sub-partnerships
	Away Team members	The Away Team attends Learning Sessions, and brings back learning to Workstreams and individual services. It also was expected to provide leadership for the EYC at a local level, particularly before the appointment of a dedicated PM.	
Operational level	Home Team members	Takes direction from the Away Team in coordinating tests of change. Participates in local Home Team events which bring members together around focussed topics.	Home Team maintains existing organisational accountabilities, participation within EYC on a purely consensual basis

Case study specification and research approach

The previous section has described the overlapping and multi-polar structure of the EYC and described its two units of analysis: the strategic-level and the operational-level. This section builds on this conceptual analysis to complete the specification of the case study boundaries, describing the case framing and site selection process.

Multiple embedded case study approach

For positivist methodologists like Yin (2009), multiple cases represent multiple experiments (rather than merely multiple observations), which provide a form of external validity. From a realist perspective however, the value of multiple cases is not

the repeated testing of hypotheses through individual experiments, but in developing a richer contextualised understanding of shared phenomena through cross-case comparison (Easton 2010; Tsoukas 1989). This allows the naturalistic exploration of phenomena across diverse contexts, allowing the isolation of contextual dependencies and the strengthening of causal theories. Multiple cases also hedge against risk, for example of CPP EYC programmes ceasing, or researcher access being overturned, risks which were inherent given overall CPP control over EYC operations. Multiple cases however entail proportionately more time and resources than do single cases, so three cases were judged to provide a suitable trade-off between breadth and depth of analysis.

Splitting the EYC into 'strategic' and 'operational' levels yields two units of analysis, with multiple TOCs operating within a CPP context. It was expected also that these two units were interrelated: the character of TOCs emerging from the Home Team was likely to influence the strategic direction taken, while this strategic direction in turn would influence Home Team recruitment and TOC development strategies. Yin's (2009) embedded case study design involves looking at one or more sub-units within a larger 'case', allowing more focussed analysis of components of interest and increasing the scope for comparative analysis within cases, across cases, and across all subunits (Baxter and Jack 2008). Eleven TOCs were chosen across CPPs, based on a theoretical sampling methodology described in greater depth in the next section.

Sampling and site access

Strategic level

Positivists base sampling decisions on how to maximise representativeness and minimise the interference of extraneous variables on the relationships of interest. Accordingly, large random samples are the gold standard of sampling methodologies. Qualitative research instead tends to adopt a purposive or theoretical sampling strategy targeting the most 'information rich' cases (Curry et al. 2009; Patton 2002) as

exemplars for a particular phenomenon of interest (Stake 1995). Case studies also aim to generalise through theoretical insight, rather than statistical inference to develop generalisable theory (Yin 2009). The sampling strategy employed therefore sought to eliminate information-poor cases (cases featuring highly irregular context, or which faced comparatively low levels of complexity), and maximise research depth in fieldwork. Four purposive filters were used to ascertain a shortlist of potential sites:

(1) Larger CPPs were considered to better embody the challenges of multi-agency working than smaller CPPs (relevant to RQ 3) while providing a crude measure of complexity (number of agencies and actors involved). Therefore it was decided to exclude smaller local authority areas (population <100,000) to ensure the challenge of complex multi-agency working was maximised.

(2) An additional signifier of the challenge of complexity was the difficulty of achieving Stretch Aims to maximise the need for transformation (IHI 2003). In the absence of baseline Stretch Aim data, the 2012 Scottish Index of Multiple Deprivation (Scottish Neighbourhood Statistics 2012) was used as a proxy to finalise a theoretical sample. The 12 CPPs with over 10% of their datazones within the 15% most deprived in Scotland were included in the final shortlist.

(3) Highland CPP was excluded owing to the unique challenges to service integration presented by its rurality, which are highly atypical among Scottish local authorities.

(4) All remaining CPP websites were investigated to ascertain that EYC had been set up according to Scottish Government guidance. One CPP appeared to have integrated their EYC very differently, and was excluded from the shortlist.

These four filters resulted in a shortlist of seven CPP areas[3], which were all judged to be information-rich and epitomise the challenges which QICs face in operating in

---

[3] In line with with the procedures followed to protect anonymity described in the following section, this shortlist is not explicitly reported.

complex social service systems, an important factor to RQ 3. To further reduce the sample to the final sample, a randomisation process was considered, but was judged to add little to the study's validity following a theoretical sampling approach. Instead, a convenience sampling method was adopted, excluding CPPs which were located outwith a 100 mile radius to minimise resource expenditure and improve accessibility. Convenience sampling increases the risk of selecting information-poor cases (Patton 2002), however this criticism is less valid following a robust purposive sampling process. Relative to randomisation, this approach allowed the resource budget to stretch further and therefore improve depth of analysis, a crucial aspect of reliability in critical realist and other post-positivist case study research (Easton 2010; Stake 1995).

PMs in remaining CPPs were sent an email introducing the researcher and the study, and requesting a meeting or phone call to discuss participation. PMs were recognised as gatekeepers to vital study participants at the strategic and operational levels, therefore the purpose of these initial meetings were to discuss access, assess appetite for the research, and ensure the case was appropriately structured to generate meaningful data of relevance to the Research Aim and RQs. Only three PMs responded agreeably to this initial email, with two phone calls taking place and one face-to-face meeting. The lack of response from the other two sites was taken to reflect low appetite for engaging with the study, and the final 3 CPPs were finalised following initial meetings.

Operational level

The sampling methodology for the embedded case study units was also theoretically informed, and conformed to a 'maximising variation' (Patton 2002) sample in order to capture the operation of the methodology across various contexts and stages of development. It was determined that to address study aims, TOCs would be chosen which were:

(1) Active, to allow potential for further developments during the research process.

(2) Representative of a diverse range of sectors, to address the multi-agency context of outcomes-focussed work.

(3) Representative of all stages of the TOC development process (framing and initiation, testing and improvement, and scaling and spreading).

Owing to the lack of a systematic TOC database at time of site access, the researcher was dependent on PM insight for both knowledge of and access to TOC case studies. This created the potential for selection bias since PMs might choose to put forward 'exemplar' TOCs, which could in turn bias the sample towards those contexts most conducive to the EYC's methodology. While this risk was mitigated by presentation of clear selection guidelines to PMs, it was decided to maximise variation in context to further minimise this risk. Following first interview, PMs in CPPs 1 and 3 were asked to shortlist all TOCs which met the above criteria. PMs in CPPs 1 and 3 outlined fewer than 10 fully active TOCs apiece, and conversation about points (2) and (3), allowed the identification of TOCs without significant PM influence. In TOC 2, where far more TOCs were claimed as active, there was greater potential for PM selectivity. Following finalisation of the TOC sample in CPPs 1 and 3, discussion with the PM in CPP 2 about TOC selection was then based on finding complementary contextual settings, rather than performance.

Following this, PMs connected with TOC leaders to discuss access. All initially selected TOCs in CPPs 1 and 3 were agreeable to participate, however it was not possible to access a school-based TOC in CPP 2, creating the absence of an educational context within the sample. Following discussion with the PM, a secondary school TOC,

ostensibly associated with RAFA[4], however pursuing a TOC project with the same CPP-level support and methodology, was selected to fill this gap, and it was decided to take on another TOC (TOC 5) at the same level of development to ensure comparability across cases. It was also decided to follow TOC 4 in CPP 1, which had adopted and then abandoned the MFI to form a comparator case and a de facto counterfactual for where the MFI was not utilised. The final TOC sample is outlined in Table 7 below.

Table 7. Test of Change case study sample characteristics

TOC	CPP Area	Service Context	Stage of development
1 Rainbow Project	1	Allied Health practitioners	Advanced
2 Stay and Play	1	Nursery / community services	Intermediate
3 What Matters to You	1	Strategic planning	Beginning
4 Family Pool Time	1	Leisure / family activities	Abandoned MFI
5 Child Development Centre	2	Parental attachment and development	Intermediate
6 Family-based Model	2	Parenting skills	Advanced

---

4 This CPP was pursuant of an integrated approach to its improvement programmes, with substantial overlap between the EYC and RAFA.

7 Vulnerable Family Engagement	2	Physical therapy services	Beginning
8 School Literacy Project	2	Schools	RAFA - Intermediate
9 Income Maximisation in Pregnancy	3	Maternity services	Advanced
10 Engaging Dads	3	Father engagement	Beginning
11 Staying Put	3	Housing and homelessness	Intermediate

Methods and research process

Case studies necessarily incorporate different forms of qualitative data (Merriam 1998; Stake 1995). While quantitative data were incorporated as context (including a CPP-level survey which the researcher co-developed, summarised in Appendix 4, and TOC documentation), a quantitative approach could not by itself produce deep insight into complex systems change (Byrne and Callaghan 2013). Yin (2009) lists six forms of data appropriate for case studies: documentation, archival records, interviews, direct observation, participant observation, and physical artefacts. The research strategy has focussed on the first four methods, which are described in the following section. Data collection was conducted between April 2014 and February 2016 (with one observation of an EYC Learning Session in November 2016). In line with Haynes (2008), the considerable length of the data collection process allowed the effects of key events to be tracked over a considerable period of time.

Interviews

Semi-structured interviews were the primary method of inquiry, generating in-depth perceptive data relative to all research questions. Interviews can range from the entirely unstructured and atheoretical, for example within Grounded Theory (Glaser and Strauss 1967), to structured verbal questionnaires giving no scope for interviewee voice. Semi-structured interviews sit between these two approaches, guided by an ordered set of topics or open questions but permitting focussed exploration of participant responses. Semi-structured interviews enabled focussed inquiry into thesis RQs while allowing the flexibility to respond and probe important issues as they emerged (Kvale 1996).

Semi-structured interviews are noted for their strength in gathering interpretive data (Mason 2002), being able to capture the meaning and value that participants attach to relevant events (Weiss 1995). They are also useful instruments in surfacing espoused rationales behind participants' own behaviours, capturing explanations of 'why [participants] feel, act and believe as they do' (Lewis and Ritchie 2003, p.253). Interviewees were not therefore considered 'vessels of answers' (Gubrium and Holstein 2002), but sources of perceptions, a view more compatible with the constructivist epistemology of critical realism.

Two cohorts of interviews were carried out at the CPP-strategic and operational levels. Interview participants were strategic level representatives including PMs, past Workstream leads, Away Team members, and operational-level TOC leaders[5]. Each interview was guided by a bespoke topic guide, which was created based on RQs,

---

[5] In one case (TOC 4) which had abandoned use of the MFI, it was decided to interview an individual involved with the initiative during the time that the project was more significantly with the EYC.

participant position, and any important aspects of the EYC's operation which were uncertain.

Strategic level

Interviews with strategic-level participants aimed primarily to generate data relevant to RQs 1 and 3. These interviews were more explorative in nature than the operational-level interviews, relying heavily on probing questions where further clarity was needed (Kvale 1996). Participants included:

- Three formal interviews were arranged with PMs at the beginning of data collection, and at three further interviews were carried out at the end of primary data collection (December 2015 - February 2016), allowing changes in views, perceptions and strategies to be assessed. Topic guides (Appendices B and C) were focussed on the perceptions of the significance of key events in the EYC's development, intentions behind strategic development, perceptions of the utility of National-level infrastructure, and perceptions of the suitability of the EYC's methodology and infrastructure.

- Four subsequent interviews with current and former Workstream members of both Home and Away Teams were arranged to supplement PM interviews, to capture different perspectives of shared phenomena, and fill in understanding of the EYC's functioning which observations and document review could not elucidate. Interviewees included one past Workstream lead (to capture perceptions on how defunct Workstreams used to function) current Workstream members in CPP 2 (to capture perceptions on their effectiveness), and 2 other members of the Home and Away teams in CPPs 1 and 3.

- Lastly, the degree of integration with the Locality Model initiative in CPP 1 (a CPP-led whole-system approach to service reform implemented contemporaneously with the EYC) and the realisation that it could serve as an important comparator case (Weiss 1995) regarding RQ 2, led to an expansion of the data collection approach to incorporate four interviews with core staff:

two community workers, and one senior officer (with the senior officer interviewed twice over a one year period).

Operational level

Operational-level interviews investigated the learning, innovation and sharing process and thus contributed to RQs 1, 2 and 3. At time of site access, little information was available as to how TOCs operated in practice, and how the testing and scaling process occurred outwith primary care settings. For this reason an arrangement was made with the PM in CPP 2 to conduct preliminary 'pilot' interviews with TOC leaders at a Home Event in October 2014.

Pilots are often adopted as 'trial runs' to stress test data collection instruments (Polit et al. 2001), however can also serve to refine the content and procedure of data collection (Yin 2009). 20-25 minute long Interviews were conducted with four individuals leading active TOCs, which were broadly unstructured and aimed to explore participant's experiences engaging with the EYC and using the MFI to generate, test and scale improvement ideas. Following these, a 45 minute-long focus group was conducted with six individuals leading active TOCs to further probe experiences, challenges and benefits which the structure and methodology of the EYC had created for them.

This pilot phase surfaced a number of important issues which were pivotal in shaping research strategy for data collection at the operational level:

- TOC leaders were struggling with reconciling the need for quantitative measurement with the nature of their services, and the methodology was often significantly adapted in practice.
- Consideration of what was important to be improved was often side-lined as practitioners moved straight into testing.

- In two of the interviews and during the focus group, it surfaced that service users themselves had been a key part of the innovation process.

These findings helped refine the topic guide used for operational-level interviews, and in particular ensured co-production became a key focus. Operational-level participants were asked to send all available project documentation (progress reports, run charts or other internal documents) by email for review prior to interview. From this documentation, tailored topic guides were developed (an anonymised sample can be found in Appendix D), however interviews were more structured than the strategic level in order to facilitate cross-case comparison. Following interview, updates about TOC case study progress were sought through the PM[6]. Finally, each TOC project was written up as a 1000-1500 word summary detailing their creation, development and improvement progress to facilitate comparison during data analysis.[7]

Across strategic and operational levels, 28 interviews were conducted with 25 individuals, along with one focus group with six participants. At the strategic level, 13 interviews were conducted with 10 individuals. This included 4 individuals representing (current or former) Workstream members, 6 interviews with the three PMs, and 3 interviews with four individuals at the Locality Model. At the operational level, fifteen interviews were conducted. This included 4 interviews during the pilot phase, and 11 during TOC case studies. Interviews ranged in length from an average of 45 minutes with Workstream members at the strategic level, an average of one hour for TOC leaders, and an average of 90 across all PM interviews. Pilot interviews were much shorter, averaging 20-25 minutes each. All interviews were audio recorded,

---

[6] No contact could be made in the case of TOC 8.

[7] These summaries are not included in Appendices because of their size (Approximately 15,000 words) but are available from the researcher on request.

summarised, then later professionally transcribed and inputted into NVivo 10 for analysis.

Observations

A complementary mode of data collection was non-participatory direct observation of Workstream meetings, Home Team events and one Learning Session. Observations are suited for data collection within a naturalistic setting (Curry et al. 2009), and from a realist perspective are particularly appropriate as complements to interviews since they permit access to data independently of their perception by others (Tjora 2006).

It was initially devised to observe Workstream meetings in all CPPs, where direct observation could access the processes of innovation and collaboration crucial to the EYC's viability as a learning system. However on site access Workstreams had remained functional only in CPP 2. Accordingly, it was arranged to observe monthly Workstream meetings in CPP 2 across the latter half of 2014, and to revisit the potential for observations in the other CPPs. In CPP 1, where much EYC infrastructure had collapsed, it was arranged instead to observe the PM in contact with TOCs through direct project development sessions, while in CPP 3 it was arranged to observe some of its own thematic infrastructure, including a thematic grouping leading improvement projects within nurseries, and a high-level ICS board meeting to which the EYC directly reported. In addition, it was arranged to attend Home Events – half or full-day events bringing together CPP Home Teams to share learning and deliver training – across all CPPs. Finally, observations were taken at Learning Session 8, attended at the end of the study, in order to assess the interaction between EYC Away Teams and how content and storyboards were engaged with.

It was decided to adopt a loosely structured observation protocol (Mulhall 2003) to document instances of innovation and collaboration in Workstreams, which can be found in Appendix E. This protocol distinguished between 'events' – for instance instances of collaboration, decisions changing, disagreements, or actions being agreed, and researcher thoughts regarding the explanations of this behaviour. Observational data from other the fora attended including Home Events and the national Learning Session were recorded using an unstructured approach using a field diary. In total, eight Workstream meetings (four each of Workstream 1 and Workstream 3) from August 2014 – December 2014 were observed, four Home Events were attended across 2014 and 2015 (two each in CPPs 1 and 2), one ICS board meeting and one thematic working group meeting were observed in CPP 3, and one national Learning Session was observed in November 2016.

Document review

A final data collection instrument was an extensive document review comprising both strategic and operational level artefacts. Document review is an unobtrusive and low-cost process which is often used to supplement data collected through interviews and observations in case studies (Bowen 2009, p.31). In the perspective of Merriam (1998, p.118), documents 'uncover meaning, develop understanding, and discover insights relevant to the research problem'. While 'found rather than made' (Jensen 2002, p.243) documents are created to shape as well as reflect their world, particularly within a policy context. In lieu of this active role, Bowen (2009, p.30) summarises the potential contributions of document review as: providing contextual information, priming the empirical analysis, adding supplementary data, providing a means of tracking change and development, and corroborating findings from other sources of data.

Documents were sought from the strategic-national level, the strategic-CPP level, and the operational level. At the strategic-national level, policy documents including Early

Years Taskforce minutes, evaluations (Scottish Government 2014) and other position papers, and Parliamentary evidence (Scottish Parliament 2014) were analysed to improve understanding of the historical roots of key events and understand official rationales for strategic decisions (Bowen 2009). Documents and artefacts from Learning Sessions including the substantial video catalogue documented from all Learning Sessions[8], and presentations, picture libraries, substantial delegate packs, and other artefacts documented on the Scottish Government's website[9] were particularly important in capturing how the EYC was presented to Away Teams, as well as providing detailed documentary evidence for the historical development of the EYC.

At the CPP-strategic level, documents provided crucial contextual data including organisational charts, Children's Services Plans SOAs and children's services inspection data, and provided data about internal EYC strategic management including project and progress reports, newsletters and communication and policy papers. These documents provided a rich timeline of local developments which could be compared with contemporaneous Scottish Government directives.

At the operational level, documents primed empirical analysis, allowing interview guides to be sharpened to focus on key points in project development. Project data was sought also for other TOCs projects to corroborate the experience of case study TOCs. Substantial documentation detailing project formation, achievements and MFI-related evidence was procured for over 20 of the best developed TOCs across the CPP case studies, which comprised more than one third of the total active TOC population across case study CPPs.

---

8 A video catalogue can accessed at https://vimeo.com/user8437545

9 http://www.gov.scot/Topics/People/Young-People/early-years/early-years-collaborative/learning-sessions

Data analysis

The research generated a considerable amount of data across many different formats, which was anticipated to present a substantial data management and analysis challenge. Accordingly, data from observational protocols, interview summaries and transcripts, and documents and artefacts, were inputted as they were collected into NVivo 10, a qualitative analysis software chosen for its strengths in data management and thematic analysis. NVivo's powerful coding and presentation capabilities allowed analysis by theme, code or source across the whole concourse of data.

Computer-aided analysis can create a tension between providing detailed, context-rich data of whole transcripts, and unhelpfully stripping context away through reductive thematic coding (Lewis and Ritchie 2003) – a particular problem if actions are considered to depend on context for meaning (Lincoln and Guba 1985). The researcher in a critical realist tradition must then be careful to use software to 'give [data] shape without doing violence to them (Richards and Richards 1994, p.83). The researcher went back and forth from data tables and thematic categories to whole transcripts over the data collection and analysis process to make sense of data while ensuring contextual dependencies were not erased through reduction.

As is standard in case study research (Patton 2002; Miles and Huberman 1994), data collection and analysis were intermeshed, with new data leading to the amendment of codes and themes, and further thematic analysis surfacing new questions for interviews and directions for document review (Weiss 1995). Following each interview detailed written summaries were produced based on the audio recording. These, along with contextual notes about the interview setting, context and researcher observations, were inputted into NVivo, where descriptive codes were attached. Notes were made relating these to thesis RQs, and codes also informed the design of interview protocols on a rolling basis.

Coding and interpretation

Firstly, descriptive open coding (Miles and Huberman 1994) was adopted to analyse the data inductively. This assessed regularities across common themes – for example relating to the value of the MFI, barriers to sharing information, types of improvements realised – which over time made commonalities clear across interview summaries, observations and documents. Throughout the data analysis process, codes were revised and re-ordered. At the operational level for instance, the code 'barriers' was used to denote difficulties in applying the MFI, however on closer analysis this code was divisible into barriers associated with 'institutional environment' and 'measurement difficulties', which could further differentiated to aspects of TOC context.

Data analysis was also crucial in beginning to link together themes from the strategic and operational levels to construct cohesive responses to RQs 1 and 3. Codes across interview transcripts were grouped together into higher-order themes (Miles and Huberman 1994): the importance attached to data and measurement (coded 'measurement focus') at the strategic-level could be joined with practitioner desire to demonstrate quantitative improvements to attract project support (coded 'selling') into a higher cultural-related code 'managerialism' which defined in clearer terms a key logic of EYC strategic management in practice, and also resonated with Quality Improvement literature.

The more standardised format of operational level interviews enabled more focussed cross-case analysis of responses and direct comparison of contextual factors using 'grid' analysis comparing contextual factors with response type. This was crucial in enabling the contextual sensitivity needed to adequately tackle RQ 3. Integrating data analysis and collection allowed a high degree of reflexivity in the fieldwork process aligned with Lewis and Ritchie's (2003, p.49) observation that, 'the relationship between design, data and theory is a multi-directional one'.

A transcription service was commissioned for each significant strategic and operational-level interview in January 2016, after which transcriptions replaced their summaries within the NVivo project file. The introduction of transcripts provided the basis for a more forensic approach to coding following data collection, attached more concretely to 'chunks' of text (Miles and Huberman 1994). This enabled more careful pattern analysis, and the further refining of codes developed during the data collection process. This was important in moving data analysis towards a retroductive process as themes became more distinct and patterns emerged more clearly in the data (Sayer 1984). This also enabled the coding scheme to become more selective based on aspects relating to the theoretical framework for outcomes management. The data could then be linked more systematically to the instances and effects of self-organisation, learning, and attractors, strengthening the link between empirical data and the propositions of the theoretical framework (Yin 2009).

Triangulation

Since it assumes all sources of knowledge are fallible, critical realism attaches a special weight to triangulation between data sources (Jick 1979). Since validity is compromised by overreliance on any particular source, where multiple forms of data converge, we can be better assured of causal relationships. Where they diverge however, the intention is not to attempt to neutralise bias or correct errors in a positivistic sense, but to represent multiple perspectives of a shared phenomenon (Baxter and Jack 2008). Thus triangulation across data sources is used to substantiate causal claims and achieve a deeper and more reliable representation of reality.

Ethics

Prior to negotiating site access, ethical approval was obtained through the School of Applied Social Sciences Ethics Committee at Stirling University in April 2014. The

approval sheet is included in Appendix A. In addition, the study was required to pass ethical review in CPPs 2 and 3[10].

The study's ethical considerations, particularly its treatment of informed consent and confidentiality, were informed by the ESRC's (2010) Framework for Research Ethics. This identified two linked risks associated with participation, surrounding anonymity and the potential for reputational damage.

Anonymity

To encourage honesty and openness in participants' responses, the thesis sought to protect participants' anonymity as far was feasible, which included concealing the identity case study sites. However, it was also important to capture contextual features of case study sites in sufficient depth, which inevitably created the potential that the identity of case study sites (and therefore the identities of key individuals within them) could be deduced, particularly by those with insider knowledge.

Agreement was reached with PMs on a form of limited anonymity. Detailed contextual and historical context would be included, however particular distinguishing details (specific demographic or geographic details or historical events for instance) would be obscured. This preserved a limited but meaningful level of anonymity which prevented casual identification of case study sites, although could not guarantee the prevention of identification by insiders. This risk was agreeable to PMs, and was also noted on interviewee participant information sheets.

Data protection principles as outlined in the UK Data Protection Act 1998 were followed in the storage of data, with only the researcher having access to audio files

---

[10] The PM did not deem it necessary to proceed through internal ethical review in this CPP.

and transcripts stored on a password-protected computer to which only the researcher had access. In the reproduction of data, specific job titles and distinguishing information were erased, and quotes used in reporting data were screened for any information which could compromise anonymity.

Reputational damage

The second related risk was reputational damage (and the fear of this) resulting from negative or critical accounts of performance. It was emphasised that the research did not seek to make judgements about the implementation of the EYC (e.g. methodological fidelity to the MFI), or to compare performance across CPPs, but to explore its functioning in a naturalistic context. This risk was also discussed with PMs at outset, however being accustomed to evaluation and critical reflection in their working roles, all participants were happy to proceed. To mitigate against this risk, all participants were informed of their right to pause the recording, redact information divulged even following interview completion verbally and on information sheets, although none did so.

Reflexivity

The researcher's previous career experience has sparked interest in the topic and entailed direct exposure to key problems detailed in the thesis, including the paradoxical effects of narrow performance management processes in the public sector. While every attempt has been made to systematise the research process – for instance in clarifying clear case study selection criteria – the research design, content and procedure have undoubtedly been influenced by the researcher's personal viewpoints and direct experience.

It is also clear that researcher interests shaped the evolution of the research project. As a clear example, the research project did not necessarily need to explore co-production in the context of the EYC to the extent it did, however the researcher's

knowledge and interest in the literature made clear the potential contribution of the study findings, and clarified how it was linked to the research aim.

Social science is now mature enough to recognise that such 'bias' is an evitable part of the research process, and that a second-order awareness or reflexivity is necessary to understand and challenge the filters through which the researcher engages with the research process and interprets findings. To aid in this reflexive process, detailed notes were taken in a journal at all site visits and interviews. While these notes were not inputted into the main Nvivo project file, they provided an aid for reflection during the research process. As the research design explains further, the rolling basis of data collection and analysis ensured that these reflections became incorporated into the data collection process, and informed decisions made for further interviews or site visits.

Summing up

This section has described the data collection and analysis procedures which constituted the fieldwork process, which is summarised in Figure 7 below.

Figure 7. Fieldwork process map

Ethical Approval: **April 2014**

First contact with PMs: **April – July 2014**

Operational level pilots: **September 2014**

First PM interviews: **May – August 2014**

TOC case studies: **September 2014 – October 2015**

Observations of Workstream meetings: **August – December 2015**

Final PM interviews: **November**
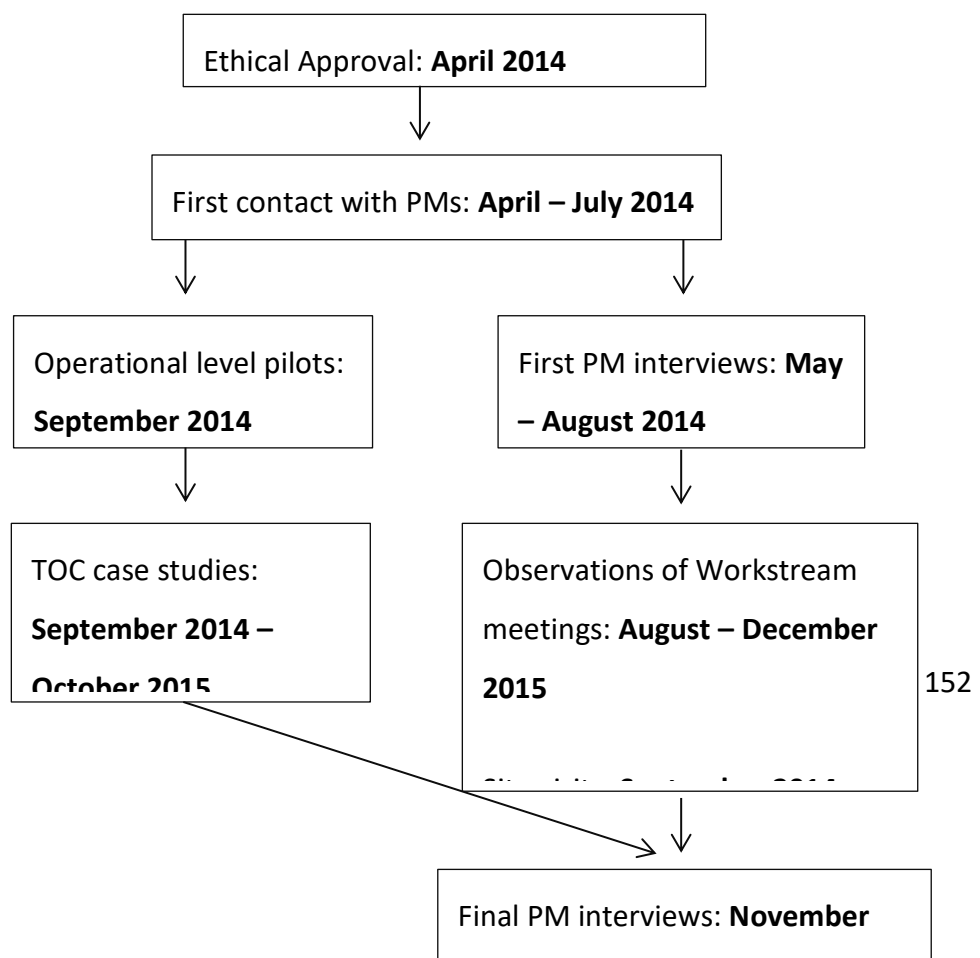
Taken together the fieldwork process gathered a substantial pool of data from which responses could be generated to RQs 1-3. The scale of the data collected required two substantial empirical chapters to report. Chapter 5 introduces strategic level-data, providing a descriptive and comparative analysis of CPP case studies. Chapter 6 reports the operational-level data concerning embedded TOC case studies. Following this, the data are re-interpreted relative to the study's three research questions in the analytical chapters 7 and 8.

Conclusion

This chapter has provided a transparent and thorough description of the research design and its relevance to the thesis aim. It has situated the thesis' philosophical position within a critical realist paradigm, and described in detail the design of its multiple embedded case study approach. The chapter has clarified the origin and purpose of the three Research Questions, which together with the last two Research Objectives structure the empirical analysis in Chapters 5-8. The following two chapters provide an in-depth analysis of the CPP case studies and TOC embedded case studies respectively, which enables the theoretically-engaged analysis in Chapters 7 and 8 to tackle the three research questions directly. Chapter 9 then brings all of this empirical analysis together in a summative reflection on the thesis aim, and reflects on how this can inform future research.

Chapter 5: development and progress in the Early Years Collaborative: insight from the strategic level

Introduction

This chapter introduces the reader to the three case study areas, and presents a descriptive and comparative account of the development of the EYC both nationally and within case study CPPs over the first three years of its development (October 2012 – December 2015). This chapter contributes to the research aim in two ways. Firstly, it draws on empirical data from the strategic-level of the EYC to impart a detailed account of progress and strategic trajectory within each CPP, allowing the reader to access later analytical chapters where the principal insights of the thesis are generated. Secondly, it contributes substantially to RO3 by undertaking a critical comparative analysis of EYC progress across CPP case studies to understand how improvement was pursued and achieved.

The development of the Early Years Collaborative: the national case

Policy Context

Children's services in Scotland since devolution have been driven by two overarching and complementary policy frameworks. Firstly, the publication of the Getting it Right for Every Child (GIRFEC) framework (Scottish Government 2008) in September 2008 positioned the protection and promotion of children's wellbeing as the central pillar for policy, practice and legislative reform. Secondly, the Scottish Government's (2008) Early Years Framework developed with the Coalition for Scottish Local Authorities and published in December 2008 committed central and local government to a strategic change agenda focussed on partnership working, local autonomy, and a shared commitment to prevention and early intervention.

In 2011 the strategic and policy coordination forum called the Early Years Taskforce was established to provide strategic leadership across the public sector and aiming to accelerate the conversion of the principles in the Early Years Framework into practice. The Early Years Taskforce published a strategic vision statement (Scottish Government 2012) which set out ambitions to:

- Deliver tangible improvement in outcomes and reduce inequalities for Scotland's vulnerable children.
- Put Scotland squarely on course to shifting the balance of public services towards early intervention and prevention by 2016.
- Sustain this change to 2018 and beyond.

To support these ambitions, two significant interventions were made. Firstly, a £274.25 million Early Years Change Fund was established in 2011, jointly funded by the Scottish Government, local authorities and the NHS. The Change Fund intended to create the financial means and shared will to convert the ambitions of the Early Years Framework to practical action and support a strategic, long-term and joined-up approach to service reform. Secondly, the EYC was introduced in October 2012 to provide a structure and shared method to facilitate the actions required to meet these ambitions.

As discussed in Chapter 3, the idea for the EYC emerged through the perceived success of the SPSP, which had adopted the BSC model to achieve substantial improvements in the prevention of hospital-related infections and patient harm (Chief Medical Officer for Scotland 2012). The idea was developed within the Early Years Quality Improvement Unit within the Scottish Government to apply the BSC model to the children's services sector in early 2012, and was presented to the Early Years Taskforce for refinement and authorisation across 2011 and 2012. The EYC was developed as an outcomes-focussed, multi-agency national quality improvement programme, delivered by the Scottish Government, Community Planning Partnerships, and the wider

children's services workforce, and launched at an Early Years Taskforce event in October 2012.

The Early Years Taskforce set up a Practice Development Team in November 2012 consisting of seven staff seconded for two years to support the Change Fund and the wider Taskforce agenda. Owing to the lack of dedicated Improvement Advisors on the introduction of the EYC however, the Practice Development Team became largely subsumed with direct operational support for the EYC (Scottish Government 2014).

In line with IHI guidance (IHI 2003), the Scottish Government set up a programme of quarterly two-day Learning Sessions, which featured plenary speakers including high profile Scottish Ministers and public service leaders. An e-learning environment (the Extranet) was also set up for participants to upload their improvement work, share learning and collaborate across physical boundaries beyond the Learning Session. Learning Sessions were delivered centrally by the Scottish Government with close collaboration with the IHI, and were chaired by the Scottish Government's Quality Unit Clinical Director. The EYC was championed as a whole by Harry Burns, then Chief Medical Officer for the Scottish NHS and a key member of the Early Years Taskforce. The development of the EYC was led by the Early Years Quality Improvement Unit, however the EYC itself had no dedicated manager until the appointment of a Head of the EYC following Learning Session 3.

Prior to its first Learning Session, the Scottish Government's Early Years Quality Improvement Unit developed three high-level 'Stretch Aims' in consultation with the Early Years Taskforce, representing age-based developmental outcomes across local populations. Stretch Aims were designed in line with BSC guidance (IHI 2003) to provide measurable system-level targets which would require fundamental system change to achieve, rather than merely incremental improvements in existing services. However Stretch Aims 2 and 3 were not systematically implemented by EYC initiation,

and thus the EYC had no baseline from which to measure success. The original Stretch Aims were:

1. 'To ensure that women experience positive pregnancies which result in the birth of more healthy babies as evidenced by a reduction of 15% in the rates of stillbirths (from 4.9 per 1,000 births in 2010 to 4.3 per 1,000 births in 2015) and infant mortality (from 3.7 per 1,000 live births in 2010 to 3.1 per 1,000 live births in 2015).

2. To ensure that 85% of all children within each Community Planning Partnership have reached all of the expected developmental milestones at the time of the child's 27-30 month child health review, by end-2016.

3. To ensure that 90% of all children within each Community Planning Partnership have reached all of the expected developmental milestones at the time the child starts primary school, by end-2017.

The launch of the EYC

In the Autumn of 2012, the Practice Development Team engaged all 32 Local Authority Chief Executives to encourage signing up to the EYC. When making contact, it was explained that the EYC would not be accompanied by additional financing and CPPs themselves would be responsible for resourcing and staffing. It was also explained that the EYC was not a statutory obligation and signing up was at each CPP's discretion. Nevertheless, all 32 CPPs responded positively and agreed to participate, which entailed a significant burden for both the Early Years Quality Improvement Unit and the Practice Development Team for operational support.

The EYC was launched on 1 October 2012 by the Scottish Government, NHS Scotland and COSLA as chairs of the Early Years Taskforce, where it was positioned as the Scottish Government's principal policy intervention to embed the recommendations of

GIRFEC (Scottish Government 2008). At the launch event the EYC was endorsed by the highest levels of government and the statutory sector at its launch event, including Minister for Children and families Aileen Campbell MSP; Minister for Local Government and Planning Derek Mackay; NHS Scotland Chief Executive (and later Vice-Chair of the IHI) Derek Feeley, NHS Scotland Chief Medical Officer Harry Burns (who appeared via video message), and Scottish Government Permanent Secretary Peter Housden.

In her plenary session, Aileen Campbell MSP described the Scottish Government's vision for the EYC to delegates:

'The Collaborative will be a multi-agency, local, quality improvement programme delivered at a national scale, taking forward the vision and priorities of the Early Years Taskforce. It will draw on learning from the highly successful Scottish Patient Safety Programme and the collaborative approach it used.'

Indeed, the EYC structure borrowed the SPSP's structure almost verbatim, and was described by the Scottish Government's Clinical Director as, 'the first national multi-agency improvement collaborative'.

Go out there and test: the first year of the Early Years Collaborative

The EYC began its programme of Learning Sessions in January 2012. Learning Sessions were designed with three aims, in line with the BSC model (IHI 2003): building will, sharing ideas and improving the execution of improvement methodology. Participants then return from Learning Sessions to their local settings to implement what they have learned – these time periods are called 'Action Periods'.

In contrast to the IHI's BSC model however, there was no bundle of evidence-based interventions or 'change package' (IHI 2003) which could be transplanted into local

settings by Away Team members. Consequently, the EYC faced the far more significant task of generating, rather than merely implementing, the evidence needed to achieve its Stretch Aims. This was recognised by the national team (Scottish Parliament 2014), and the focus on knowledge-generation featured strongly within the design of the first three Learning Sessions. Learning Session 1 was designed to capture hearts and minds, but also to train attendees in the use of the MFI and encourage its adoption when Away Teams returned to local settings. Learning Session 2 would then bring participants back after the first Action Period to take stock of learning while continuing to build QI expertise. Learning Session 3 would then focus on methods for scaling up the improvements which had by then begun to emerge and take shape. By Learning Session 4, it was hoped that good improvement ideas would have become evidenced at a large scale and within and across CPPs, broadly in line with timescales in healthcare settings and as advised by the IHI (IHI 2003).

Learning Session 1 drew around 800 participants from across all 32 CPPs. Harry Burns opened the Learning Session, positioning the EYC as the key public sector platform to address the social disadvantage at the root of health inequalities. Participants were exposed to the four Stretch Aims, and practitioners were introduced to the MFI in theory and practice (relying solely on examples from healthcare settings) at plenary and breakout sessions. CPP teams at breakout sessions jointly developed Driver Diagrams and ideas for TOCs to develop when they returned. At the end of the conference, the Clinical Director encouraged attendees, now ostensibly trained in using the MFI, to raise their hands and promise to begin testing on their return to CPP settings.

While no systematic data was kept regarding improvement activity which emerged during the first Action Periods, interviewees from the three case study CPPs all reported the emergence of large numbers of TOCs emerging once energised Away Team members returned to their local context, suggesting this initial call to action was very effective. Judging from Change Fund returns, documentary evidence, and

information from case study PMs, the majority of CPPs also followed Scottish Government guidance in developing an appropriately specified EYC Home Team and setting up its four Workstreams.

CPPs lacked operational management during the first Learning Session, with most CPP PMs taking up posts around Learning Session 3. The Practice Development Team became subsumed with the provision of improvement support, however were themselves in the process of learning improvement methodology (Scottish Government 2014). Thus TOC development during the first two Action Periods took place in the absence of dedicated and trained technical advisors, and without any dedicated strategic management. In all cases, TOCs emerging during this period appear to have been poorly focussed, to have little relevance to the format of the PDSA cycle, and to be poorly documented, with little evidence being retained of their existence or achievements in any case study CPP. The vast majority of this improvement work appears to have fizzled out, even by the time of the second learning session, three months later.

Perhaps recognising the lack of progress during the first Action Period, Harry Burns introducing Learning Session 2 in May 2013 noted, 'the aim of the Early Years Collaborative was transforming Scotland (…) But my guess is that in the past two or three months, you've begun to find is you've begun to realise it wasn't quite so easy as you felt'. Learning Session 2 explored CPP feedback from the first action period and breakout sessions exposed participants to more examples of improvement methodology in practice, though still within a healthcare context. Learning Session 2 Delegate Packs (Scottish Government 2013) signposted to quantitative data sources which practitioners could incorporate within extant improvement work. Presentations in Learning Session 3 in October 2013 focussed on methods of taking learning to scale, some of which had by then been in existence for six months.

Most CPPs appear to have appointed core EYC staff by the third Action Period. Local authority Change Fund returns and informal discussion with other EYC members showed wide variation in how the EYC was staffed. The vast majority seem to have appointed a dedicated PM to take overall responsibility, with some appointing data managers to support improvement work, and some appointing communications staff to support PMs. The PM role was also implemented with high variation across CPPs. Most PMs were new to the improvement methodology, and instead drew from a variety of backgrounds working within the CPP. CPPs positioned the PM role at different levels of seniority, partly emphasising its level of priority given to the EYC within the CPP. Some PMs were appointed on a part-time basis, and some CPPs splitting the PM role into two positions, one in charge of strategic management, the other of operations management. CPPs also accessed different sources of funding to resource posts and TOC work, with many accessing Change Funds to support individual improvement work.

By the end of the third Action Period, the EYC had fallen short of initial hopes, lacking any concrete examples of TOCs reaching significant degrees of scale or spread. Real-world examples presented at early Learning Sessions thus remained within clinical and primary care settings, which Away Team interviewees and Workstream participants felt lacked relevance. After nine months of testing across all 32 CPPs, the EYC had persisted beyond the lifespan of most other large-scale QICs, yet it still lacked the evidence base necessary to move towards its intended focus on implementation and spread.

'One year in, we really need to get moving': Learning sessions four – eight

Harry Burns stepped down as Chief Medical Officer and EYC Champion at Learning Session 4. The former head of the SPSP was appointed to Head of the Early Years Collaborative, and took on dedicated leadership of the EYC prior to Learning Session 4. Three innovations were introduced at Learning Session 4.

Firstly, the Scottish Government introduced seven 'Key Change Areas' (KCAs): Early support for pregnancy and beyond; Attachment, child development and learning; Continuity of care in transitions; 27-30 month child health review; Parenting skills and family engagement to support learning; Addressing child poverty; Health and Wellbeing, and later an eighth, 'Play', was introduced. KCAs supplemented the Stretch Aims and their associated Driver Diagrams in providing strategic focus to improvement projects; however they reflected broad thematic areas rather than the set of evidence-based interventions customary within the BSC model. KCAs were described in delegate packs (Scottish Government 2014, p.6) as 'big ticket items', which 'if done every time for every child, will result in the achievement of the stretch aims'. This issued a clear signal reinforced throughout the Learning Session that some TOCs were to be given higher priority than others.

TOCs which were clearly thematically aligned with any of the KCAs could apply to become a 'Pioneer' site. Pioneers were offered additional technical project support from the Scottish Government Improvement Advisory team in return for a commitment to report monthly and engage in *Webex* training conferences. An open call at Learning Session 4 enlisted 30 interested TOC leaders, with 40 pioneer sites being operational across 17 CPPs by Learning Session 5.

Finally, the Scottish Government aimed to significantly expand improvement capacity. Two three-day 'Improvement Bootcamps' developed by the IHI were delivered in Glasgow. Bootcamps invited practitioners with improvement ideas to a focussed QI training programme which aimed to convert improvement ideas into actionable TOCs. Bootcamps were held in June and October 2013 and attracted 160 attendees, of which 60 were offered more advanced Bootcamp training in June and October 2014. Alongside this, a more intensive improvement programme was offered to recruit accredited 15 Improvement Advisors (IAs) within the Scottish Government through an intensive training programme delivered by the IHI and Healthcare Improvement

Scotland. Once trained, these IAs would replace the Practice Development Team in supporting CPPs with improvement projects.

Taken together, these innovations represented a drive for better focus, a message communicated explicitly during Learning Session 4, with the new Head warned at plenary that 'one year in there's a danger that unless we now focus on the things that are going to make a difference (…) won't spread out' a sentiment echoed in *the* later mid-term evaluation (Scottish Government 2014). Later Learning Sessions would continue this sense of focus, exhibiting a shift in emphasis from building will to a technical focus on 'execution'.

In 2014 the Scottish Government instituted a performance monitoring framework to document progress of its TOCs. PMs were asked in late 2014 to report on active TOCs using a standardised template, which rated level of development from intent to participate to 'significant improvement' (see TOC summary statistics in Appendix G). Returns collected in April and August 2015, reported the existence of 400 and 650 TOCs respectively underway across all CPPs. Despite the large number of extant TOCs, evidence of TOCs reaching a stage of sufficient scale for population-level impact remained sparse, and no evidence of significant spread from one CPP to another was noted.

From May to July 2015, the EYC began running Key Change Events alongside Learning Sessions, each of which focussed on a particular KCA. These events were contributed to by 'industry experts' in the topic area, and CPP teams which had significant relevant improvement work were invited to participate. Key Change Events were focussed events in which CPP teams reported more opportunities to share and discuss relevant work around priority issues, including notable issues such as increasing take-up of 27-30 month child health review, increasing uptake of Healthy Start Vitamins, and maximising income among pregnant women.

Quality Improvement across the lifecourse: the final year of the EYC

The Scottish Government launched its second major improvement programme, Raising Attainment for All (RAFA) in June 2014 within 12 CPPs. It has since expanded to 24 Local Authorities and over 180 schools. RAFA operated across primary and secondary schools with four Stretch Aims covering developmental outcomes until school leaving. A new Head was appointed to lead both EYC and RAFA in mid-2015, who on appointment felt that processes were overlapping and unnecessarily dislocated. The Scottish Government then began a process of rationalising and integrating its approach to supporting large-scale improvement, instituting a joint Learning Session in November 2015. The Scottish Government moved towards a full integration of the EYC and RAFA over 2015 and 2016, launching a single overarching 'Children and Young People's Improvement Collaborative' in November 2016, debadging the EYC and RAFA, to institutionalise QI from pre-birth until school leaving.

Summary of EYC development

Since the initiation of the EYC in October 2012 there has been a growing emphasis on QI methodology within the Scottish Government, as can be seen in a timeline of the EYC's development presented in Appendix I. Following the early emphasis on experimentation and unrestrained testing at Learning Session 1, there can be detected a progressive reigning in of practitioner agency and a more significant emphasis on focussed improvement, notably through the introduction of KCAs. Alongside this, there has been a continued emphasis on QI training and workforce development, reflecting a growing confidence in QI methodology. The integration of the EYC and RAFA into CYPIC and the central training of IAs within the Scottish Government are reflections of this confidence, and a signal that the Scottish Government believes the methodology to be a significant component of its future approach to public service reform.

The Community Planning Partnership experience

CPP Area 1

CPP Area 1 is a Scottish city which, compared to other CPP areas, features a high proportion of communities within the top 15% of most deprived datazones measured by SIMD, and a particularly high proportion of children living within these datazones.[11] The CPP Area is served by a large Health Board, which has operates a QI training programme for healthcare professionals.

The CPP arranges its workstreams into seven thematic areas, one of which relates directly to children's services, and features one SOA outcome dedicated to the early years. An Integrated Children's Services (ICS) structure, managed by an overarching strategic partnership is responsible for delivery against these objectives. This high-level strategic body was responsible for satisfying statutory requirements to prepare the Integrated Children's Services Plan and implementation of GIRFEC. Feeding into this strategic body were seven thematic planning groups focussed on different priority areas decided by the ICS partnership, and a strategic planning group responsible for coordination of these groups and overall resource utilisation, which the EYC on its inception reported to. Finally, an over-arching GIRFEC implementation team was in place to ensure policy coherence across ICS.

The EYC was initiated amid a significant period of flux in the CPP area's children's services landscape. A major reform programme was put in place across 2012 called here the 'Locality Model' pilot; a concerted effort to map, coordinate, integrate and understand the totality of resource spend on children's services modelled on the 'Total Place' approach to whole-system reform piloted in English local authorities (HM Treasury 2010). This initiative had the strategic support of the Chief Executive and the Head of Children's Services, and across 2012-2015 was responsible for setting up a

---

[11] Given the paucity of CPP areas which are bound entirely to cities, it is not possible to describe the precise historical, geographical or demographic details of this site further without jeopardising the limited anonymity agreed with PMs and necessary for site access.

number of community-based services (these are described in greater detail in Chapter 7).

The CPP sent a full delegation of senior officials to Learning Session 1, and set up Workstreams in line with Scottish Government guidance. Workstream leads and deputes were recruited from the Integrated Children's Services structure, and Workstreams were arranged to meet monthly. The Leadership Workstream was led by a voluntary sector leader, with the local authority's Head of Education becoming Champion.

Although the CPP signed up fully to the delivery of the EYC, strategic-level interviewees noted significant ambivalence from the council Chief Executive and senior CPP and ICS leaders. Interviews revealed scepticism about the methodological appropriateness of the MFI and confusion over the EYC's strategic direction, however an equally significant factor was the existing commitment to the home-grown Locality Model, which was already funded and commanded greater senior support. Accordingly, the EYC was somewhat low on the agenda of the CPP – particularly since the Locality Model was already linked into SOA and ICS priorities. The EYC therefore became substantially integrated with the Locality Model, viewed as a means of accelerating the projects within it, notably TOCs 3 and 4.

Strategic-level interviewees reflected that many TOCs were initiated following the first Learning Session, however that these were unanimously poor quality, both in their appreciation of the MFI, and in their overall focus, with none persisting significantly into the development of the EYC. A senior leader within the local authority noted in documentary evidence that 'we went full tilt into developing as many tests as we could. That's the wrong way to do it'.

Appointment of PM and early development

Three posts were created to take the EYC forward:

- A full-time PM, who came on board in mid-2013.
- A part-time Data Manager, appointed from the health board.
- A part-time information officer to assist with communication across the CPP.

The PM came from a Community Learning and Development (CLD) background, and brought pre-existing connections to the Locality Model, community-based services and third sector agencies within the CPP. On appointment, the PM noted that consequent to the lack of senior support, the EYC was struggling to retain interest from its Home Team membership. Workstreams failed to attract high ranking managers as leads and deputes. Key sectors, including midwifery, health visiting, and primary schools did not feature at all within Home Team membership on site access.

These factors made it difficult to sustain momentum of improvement work and consequently workstreams became increasingly dysfunctional. Strategic-level interviewees agreed this signalled that the EYC was not a strategic priority, and also prevented the creation of institutional space for improvement work to take place. The PM reported frankly at first interview that 'the Scottish Government structure has not worked for [this CPP]'. Following her appointment, the PM discussed options with the strategic planning group, whereupon it was decided to disband workstreams and bring leads and deputes into one overarching leadership group. The PM then met with the planning group every 6 weeks to report progress and discuss strategy.

In the absence of operational workstreams, the PM relied primarily on ICS strategic networks for TOC development, which was focussed heavily on the Locality Model. At least 5 of the 11 active TOCs within the August 2015 TOC return involved Locality Model services or the networks it had created to a significant extent.

A progress review surveying ICS members in September 2013 identified a lack of direction and focus, and a lack of leadership for improvement methodology within their practice. Asked 'how well is the EYC going?' respondents rated the EYC just 4 out

of 10. In response, an EYC action plan was agreed by members of the CPP's strategic planning group in November 2013 with three priorities:

- Improved direction and focus

- Agreement on the key priorities for the local authority

- Fewer, more focussed TOCs, better connected with strategic priorities.

Progress and TOC development

The PM was in regular contact with all TOCs, and took an active role in project development within all, visiting TOC leaders in regular planning sessions and providing remote technical assistance. At first interview the PM related the EYC's value 'goes far beyond the [IHI MFI]', and was located in the connections it put in place. This PM's approach to strategic management was noted by taking an active role in the facilitation and spreading of improvement projects. The PM also pushed for TOC development to take place within teams to ensure 'every TOC is a collaborative TOC' (PM, final interview). Indeed, most of the CPP's TOCs were led by teams, with only a handful led by individuals. Asked at final interview to consider the main successes of the EYC chapter following two years of development, the PM responded that 'building relationships' both among TOCs and across the wider EYC membership would promote the most lasting change.

The EYC chapter did not put forward any pioneer sites, nor did KCAs have any perceptible impact on practice, with the PM taking little notice within TOC development functions or wider strategic management. The TOC did however take advantage of Improvement Bootcamps, sending 5 individuals to the first session. On reflection, the Bootcamp was considered too short to be effective in coaching improvement. Instead, the PM had linked with the regional health board to access its tailored improvement coaching programme. This programme was much more in-depth and tailored to the individual project than Bootcamp, with a one-to-one mentoring component supporting project development over a course of several months.

Later development

16 TOCs were listed in the August 2015 TOC return, 9 of which were listed as actively testing, and 7 of which were reported as having achieved 'modest improvement' on the Scottish Government's rating scale. At first interview, the PM considered that just 'three or four' TOCs had the focus and ambition to achieve scale, drawing from past experience with TOC development paths. By final interview in early 2016, just one of these TOCs (TOC 1) had achieved significant improvements and scale, with talks in place at data collection termination to spread to three other CPP areas. Two of the other high-potential TOCs had ceased functioning when the Locality Model reached the end of its funding period in late 2015.

The CPP had received Scottish Attainment Challenge funding in mid-2015 to initiate improvement work around literacy and numeracy in over 20 primary schools and 8 nurseries. The receipt of this funding was perceived to created more of an appetite for change, and the PM leveraged her position within the CPP to initiate a considerable number of other TOCs within nurseries over 2015 and early 2016. At final interview estimated the existence of around 20 TOCs in planning or active testing, with a majority being located within nursery settings. This had also enabled the PM to resurrect and transpose learning from two previous TOCs (including TOC 3) into nursery settings.

CPP Area 2

Context and set-up

CPP 2 is subdivided into four geographical regions which vary significantly in both their rurality and level of relative deprivation. Two of these areas are characterised by significant deprivation clustered around urban centres and smaller towns associated with post-industrial decline. The CPP also features a significant component of rural poverty, with often poor transport links to its population centres.

Children's services in CPP 2 were governed by its strategic ICS partnership, which reported against one thematic outcome within the SOA. ICS partners were also responsible for developing and delivering against its 3-year Children's Services Plan, two of which were published during data collection.

Underneath the ICS partnership were a number of strategic sub-groups, including a children's rights and participation partnership, a corporate parenting grouping, and an early years planning group. The latter partnership was responsible for developing strategy and planning children's service delivery across the CPP. Reporting in turn to the early years planning group were seven locality-based centres coordinated the delivery of local children's services. An EYC project management group was set up to feed into ICS directly on appointment of dedicated staff.

In 2012, the local authority had committed a significant (over £7 million) budget over three years to support its Family-based Model, an overarching CPP policy to support an asset-based approach to service delivery based on the identification and response to local problems. The Family-based Model was aimed to complement universal services by focussing on the development of parenting skills and the strengthening of local service networks in the CPP's most deprived wards. In practice, this involved substantial local community engagement exercises and the development of new locality-based services.

Appointment of PM and early development

The opportunity to participate in the EYC was received enthusiastically by CPP leaders, who viewed the approach as highly complementary to the Family-based Model. The Away Team attracted senior figures from a wide range of pertinent organisations, including notably maternity and nursery leaders. A senior highly-regarded councillor became Champion, and the four Workstreams were dutifully set up following Learning Session 1. Workstreams attracted high-ranking officials as leads and deputes. The CPP already had an active Maternity and Children Quality Improvement Collaborative

(MCQIC), part of the SPSP, operating within its major hospital, and this was reconfigured as a joint MCQIC/EYC Workstream 1.

In the absence of dedicated PM support, a wide range of TOCs emerged in the first two Action Periods which were became seen as uniformly unfocussed and poorly executed, with some individuals submitting up to 25 individual PDSA cycles according to the PM. Unusually, both a Programme Manager and a full-time Project Manager were appointed around Learning Session 3. The Programme Manager role was absorbed by a senior figure within ICS alongside existing duties, who led on strategy and integration, while the Project Manager role was a full-time position with equivalent operational and strategic management duties to other Programme Managers. Accordingly, the acronym 'PM' in this case refers to the Project, rather than Programme Manager, however denotes the same duties and function. This PM also drew from a CLD background, and was expressive during conversations about the similarities between that discipline's tradition of reflective practice and the EYC's methodology. The EYC also drew technical support from an improvement programme manager within the Health Board.

Senior representation and management support was perceived by Home Team members to have encouraged lower-level managers and practitioners to engage, and was cited directly by several Workstream members during discussion as a rationale for engagement. The EYC project management team and PM took on a strategic push for greater awareness of and participation in the EYC across 2014 and much of 2015, although it was noted that awareness among the wider early years workforce remained low.

The CPP also initiated supported the delivery of two bespoke CPP-level Improvement Bootcamps in 2014, commissioning both IHI and local NHS improvement support. Invitations were extended throughout ICS infrastructure to non-engaging services and individuals in an attempt to expand EYC membership. One two-day Bootcamp

attended by the researcher in October 2014 led to around 10 TOCs being developed, several of which would persist to the August 2015 TOC returns.

Progress and TOC development

The functional Workstreams, comparatively high profile of the EYC, and the drive for greater awareness of the EYC resulted in a far larger number of active TOCs relative to other CPPs. Over 50 TOCs were estimated as active across 2014 and early 2015, which were later pared down to 34 by August 2015.The August 2015 Scottish Government TOC return placed the CPP within the top quintile across Scotland. The CPP could also demonstrate 5 or 6 TOCs which had achieved some form of scale, although none had achieved anything nearing CPP-wide adoption. Owing to the scale of its improvement activity, this CPP was invited to share learning at later Learning Session plenaries.

An evaluation of the EYC in June 2015 (contributed to by the researcher) confirmed high levels of engagement, however revealed a lack of methodological confidence and focus across the CPP (survey results are reported in Appendix). Improvement support and coordinating workstream meetings took up the majority of the PM's time, with support spread thinly across TOCs. Many TOCs were left largely to fend for themselves without direct support, which resulted in many fizzling out over the course of data collection, though detailed statistics of TOC termination were not maintained.

The CPP put forward two of its best developed TOCs as pioneer sites. KCAs also became part of PM practice in delivering improvement work and in reporting templates completed by TOC leaders, however the lack of direct involvement led to very limited strategic coordination. The CPP also engaged enthusiastically with Scottish Government QI training, sending seven individuals through national Bootcamp, in addition to over 20 passing through CPP-level Bootcamp. The CPP also later engaged with Key Change Improvement Advisor support for six of its better-developed TOCs. Nevertheless, well-developed TOCs remained a small minority of the total profile, with most remaining low impact, and a comparatively high rate of TOCs being abandoned.

While Bootcamps were effective in generating many new TOCs, workstreams declined in importance and efficacy. The eight meetings observed (four each of WS1 & WS4) over 2014 consisted largely of individual progress updates and discussion around shared issues. While opinion and advice were shared, this was always at a general level, with little evidence of collaborative innovation – e.g. the emergence of joint projects, confirmation of advice leading to improvement, or spread of learning from site to another.

Workstreams were also noted for their homogeneity: Workstream 1 was staffed almost entirely by NHS maternity staff, with meetings held in a regional hospital meeting room, while Workstream 3 was attended predominantly by Nursery heads. The EYC was also perceived by MCQIC members as remaining very separate to their improvement work, and descriptive observations charted escalating dissonance between MCQIC and EYC, with the EYC diminishing in importance.

Later Development

This CPP was the only case to take expansion of the Home Team and coverage across the wider early years workforce a strategic priority. However by second interview the EYC management team had published a position paper reversing this expansionist approach, consciously scaling back existing improvement work to focus on high potential TOCs with clearer links to CPP priorities. Recognising the failure of the majority of its TOCs to convert to significant improvements, and facing pressure from the early years strategic planning group, by 2015 the CPP had pursued a strategy of narrowing down its TOCs, with the PM reflecting at final interview, 'we now look at what needs to be improved rather than going out and testing everything'. This process resulted in a narrowing of TOCs to fewer than 30 by late 2015.

This CPP eventually followed the other case study sites in disbanding its Workstreams in mid-2015. Workstreams were described retrospectively by the PM as 'constraining us' and being 'separated from the landscape', contributing to a feeling of

fragmentation in children's services alongside the many other thematic groups under ICS. A mid-2015 project plan approved by the early years strategy group repealed the Workstreams, committing instead to an extended programme of QI training across the workforce. The CPP planned to invest even more heavily in QI training to substantially increase improvement training at all levels, to increase sevenfold the number of practitioners leading, mentoring and using the MFI by 2018. The CPP had also engaged with RAFA, steadily building its cohort of secondary schools. The early years strategy group agreed to proceed with the integration of MCQIC, the EYC and RAFA to create an overarching improvement collaborative in September 2015.

CPP Area 3

Context and set-up

CPP Area 3 is a large local authority area which features a large rural upland area and a number of urban burghs. A decline in manufacturing employment over the latter half of the 20th century impacted significantly on the authority area's urban centres, resulting in a considerable inequality of health, social and economic indicators (measured by 2012 Scottish Index of Multiple Deprivation between CPP districts.

Early years services formed one of the community plan objectives in the CPP's SOA, which along with obligations towards GIRFEC, were the responsibility of a strategic ICS partnership. This partnership published its six-year Children's Service plan in 2012, which placed a renewed focus on the implementation of GIRFEC's eight well-being indicators[12]. This plan also committed the ICS structure to a process of self-evaluation and continuous improvement under a shared performance management framework. This framework linked ICS to four regional coordination groups, staffed by local

---

[12] These are: safe, healthy, achieving, nurtured, active, respected, responsible and included

managers and practitioners, who took on responsibility for implementing local action plans linked to ICS ambitions.

Appointment of PM and early development

The opportunity to engage with the EYC was received enthusiastically by the CPP, with the Head of Education coming on board as Champion and the Chief Executive taking a supportive but disengaged position. The EYC was managed through a strategic coordination group, which reported directly to the wider ICS partnership. In 2013, the EYC Stretch Aims also became integrated into the CPP's reporting mechanisms for its SOA.

The CPP sent a full delegation to Learning Session 1, and implemented the four Workstreams. In common with other CPPs, a raft of unfocussed TOCs began during the first action period, the vast majority of which had disintegrated by appointment of dedicated staff around Learning Session 3. A dedicated PM was appointed in August 2013 to take the EYC forward. The PM came from a senior strategic post within community planning, with a varied background in health improvement, community care and corporate functions. No data manager or other appointments were made. This PM was appointed at a more senior level than in CPPs 1 and 2, and reported to a strategic coordination group below the ICS partnership.

By Learning Session 3 Workstreams were faltering, having lost key strategic members and becoming viewed increasingly as unproductive. The PM met collectively with workstream leads shortly after her appointment, where dissatisfaction was discovered about the poor relevance of age-based of working groups (since agencies tended to operate across age ranges), the lack of productivity of meetings, and the overlapping of work with existing ICS infrastructure.

Following this, the strategic coordination group agreed to disband the Workstream structure and seek to overlay the EYC as an approach over its thematic infrastructure.

Following a self-assessment process, four thematic partnerships had been developed: child poverty, substance misuse, parenting attachment, and developmental milestones. Consequently, the EYC's testing focus became integrated with each of these groups.

Following this shift the PM noted increased buy-in and a renewed enthusiasm among practitioners for the MFI as a tool to realise improvements against partnership goals. Following the Scottish Government's introduction of Key Change Areas in Learning Session 4, the PM described a feeling of 'vindication' over the recognition of superiority of thematic (as opposed to age-based) arrangements in supporting collaborative improvement.

Progress and TOC development

The PM at first interview considered the primary value of the MFI as 'getting practitioners to document their activity', both to improve practice locally and to facilitate better commissioning and strategic decisions. In turn, the collaborative infrastructure was devalued, with the PM stating at first interview, 'the EYC is something you do; it's not a club'. A somewhat more managerial approach was taken to TOC development relative to the other CPPs, with managerial sign-off and resourcing being put in place before beginning improvement work. There was also a substantial period of strategic planning prior to initiating PDSA cycles, where each TOC would be fitted within CPP priorities by logic modelling through Driver Diagrams.

The PM's experience with early TOCs convinced her that intensive one-to-one support was essential in progressing TOC development. Consequently, asked about involvement with TOCs, the PM responded that she was 'involved to a greater or lesser extent in all of them', estimating that 50-60% of her role consisted of hands-on technical and project support. The PM sought CPP funding for an extra part-time post to assist with improvement support for parental support and nurseries, and another short term contract was agreed with the health board for an individual to lead improvement

work around the Child Health Review. These posts each took over several TOCs, freeing up the PM to continue strategic work and network building.

The CPP also aimed to build QI skills to encourage more local leadership for improvement. A two-day CPP-wide Bootcamp was staged in June 2015, which attracted 35 individuals. This event involved Early Years practitioners who could both bring a developed improvement idea, and at least one other team member attended supporting the same project, as the PM had found team support to be crucial to TOC development. An additional QI training day was held in August 2015 with 14 attendees and plans to stage the programme quarterly. The CPP also supported continued professional development for individuals who had attended national Bootcamp to take on a mentoring role for other QI practitioners.

The CPP initiated two pioneer sites in parental engagement and pre-natal income maximisation, however despite both improvement projects persisting the pioneer status was seen to add little value to the project and was allowed to lapse by early 2015. The PM found KCAs more helpful in providing strategic focus, with these, as well as CPP priorities, were featured in strategic plans and Driver Diagrams developed for each TOC. The CPP participated prominently in relevant Key Change Events. One TOC (TOC 9) became engaged in focussed collaboration with a small number of other CPPs with similar aims, which by final interview the PM considered was effective in improving practice in other CPPs.

Later development

By August 2015 the CPP had 12 active TOCs, 3 of which had achieved a significant degree of scale within the CPP. While the focussed and intensive approach to TOC development undertaken placed a natural limit on the number of TOCs supported, the appointment of support staff had eased this burden, and enabled the degree of scale observed. When the CPP signed up to RAFA, the PM was then able to take over joint responsibility for both QICs. By study termination no specific plans were in place to

bring the programmes together explicitly, however the CPP had undertaken efforts to integrate leadership and training programmes across programmes.

The CPP had also increasingly absorbed the EYC's methodology across ICS structures. Following a Children's Services inspection, in 2015 the ICS reviewed structures and instituted an approach to embed the EYC's methodology as part of a new approach to continuous improvement across children's services. At final interview the PM related that the EYC was moving increasingly away from a separate structure towards an integrated approach: 'It's hard for me to say "these are the Early Years collaborative structures" (…) we're getting in and influencing core business as opposed to sitting in a little bubble'.

Comparative Analysis of CPP case studies

This section undertakes a comparative analysis of the case study sites described in the first part of this chapter. In a critical realist perspective, key differences illuminate important sources of variation for further analysis, while consistencies strengthen causal claims. Drawing from data collected through interviews, strategic-level documents, and site visits, and an analysis procedure rooted in open coding described in the previous chapter (Miles and Huberman 1994), Table 8 presents a comparative summary of key features across cases.

Table 8. EYC chapter profiles

		Initial EYC structure				Later EYC innovations	
	EYC leadership	Workstreams	Relation to Integrated Children's Services	Home Team	Away Team	Pioneer sites	Key Change Areas
CPP 1	PM from CLD background. Little practical senior involvement in home team	Abandoned by Learning Session 4	EYC fed into strategic planning group under ICS	Largely defunct by Learning Session 4. EYC overlapped significantly with the Locality Model, and later with Attainment Challenge nurseries	Active	Never adopted	Never adopted
CPP 2	Led by Project Manager (PM) from CLD background, with strategic leadership from Programme Manager. Senior involvement in all workstreams	Abandoned in June 2015	Programme Manager leads strategic planning group, with EYC represented on ICS board	Active, with regular events throughout 2014 and 2015, and large-scale QI training embedded within workforce development strategy	Active	2 pioneer sites active, both receiving Scottish Government IA support	TOCs report to PM on KCAs, however not integrated with strategic management

| CPP 3 | Led by senior-level Programme Manager from strategic planning and healthcare background. Senior manager involvement directly in TOC support | Abandoned by Learning Session 4, replaced with existing thematic groups | EYC forms basis of an Early Years Coordination group, directly beneth ICS | Semi-active, with occasional home team events, training and existing partnerships subsumed within EYC membership | Active | 2 active pioneer sites, later repealed and limited engagement with central team | Thematic areas pre-agreed in local authority; some overlap with KCAs. |

TOC development across CPPs (August 2015)

	Active or planned TOCs*	TOCs with active testing*	TOCs having achieved improvement*	TOCs having 'scaled up' in activity	TOCs spreading across borders
CPP 1	16	9	7	1	0 (1 in talks)
CPP 2	34	33	31	4	0

| CPP 3 | 18 | 12 | 6 | 3 | 0 (1 indirectly) |

* 'Active' and 'improvement' measures are based on self-reported Scottish Government TOC returns. Measures of scale and spread are not reported within this scale -

'scaled up' refers to testing which has moved to at least one other site within the CPP.

Experience with EYC infrastructure

CPPs were admitted substantial autonomy over local development, allowing the EYC to reflect the varied characteristics of the CPPs in which they were embedded. Judging from Change Fund returns and CPP websites, the vast majority of CPPs appeared to follow Scottish Government's structure of a Home and Away Team and set up the four Workstreams. All CPPs appointed at least a PM to provide operational leadership, with many appointing data managers or other staff. Differences between CPPs became particularly marked on appointment of PMs however, at which point local autonomy began to influence development trajectories much more substantially.

Collapse of the home teams

While all CPPs maintained an active Away Team throughout the EYC's lifespan, what is notable across CPPs is the complete collapse of Workstreams. Both CPPs 1 and 3 dismantled their workstreams immediately following PM appointment, finding them unproductive and duplicating existing work. Observation of ten of CPP 2's workstream meetings corroborates this, with no instance of TOCs spreading, being jointly developed, or instances of advice leading to tangible changes being observed. In June 2015, even CPP 2 which enjoyed high levels of senior support followed suit, disbanding the workstreams and focussing instead on broad-based QI training across children's services. CPP 3's PM was especially critical of Workstreams viewing the EYC's value solely within its methodology. Following a similar path, CPP 2's later approach integrating MCQIC, RAFA and the EYC into an overarching collaborative recognised the failure of its dedicated collaborative structures. In the cases of CPPs 2 and 3 in particular, the function of the EYC became integrated with workforce development rather than a means of sharing and learning across borders.

The priority of CPP goals

In every case, CPP-level priorities formulated within ICS partnerships were observed to shape the EYC's development far more significantly than Stretch Aims or KCAs. No PM

reported Stretch Aims or their associated Driver Diagrams as significantly influencing decisions made about planning, strategic management. Resistance was detected across all CPPs (particularly CPP 3) over the imposition of external goals (Stretch Aims and KCAs), since priorities had already been agreed through active Children's Services Plans. All CPPs were undergoing substantial reform efforts in children's services before the EYC was instituted. The EYC became integrated with CPP 1's Locality Model initiative, CPP 2's Family-based Model, and CPP 3's thematic goals and infrastructure. When Workstreams collapsed, the EYC became more dependent on these initiatives to help generate and maintain TOCs.

Limited Engagement with EYC innovations

Case study CPPs also exhibited limited engagement with EYC innovations introduced at Learning Session 4. CPPs 2 and 3 initiated two pioneer sites each, drawn by hopes it would generate increased exposure to other CPPs and the Scottish Government. CPP 2 was also able to attract Improvement Advisor support for six of its better-developed TOCs, and cited having previous contact with Scottish Government IAs as important attracting this support. CPP 3 however allowed pioneer status to lapse, finding the support provided not conducive to project advancement.

All PMs reported making some effort to incorporate KCAs in TOC development however KCAs were so broad in theme that most TOCs were already thematically aligned. CPP 3's PM linked each TOC to strategic CPP outcomes and KCAs through Driver Diagrams before testing, however KCAs were not considered to influence TOC development in practice by any PM or TOC leader interviewed. At an ICS strategy meeting in CPP 3 observed by the researcher, KCAs were opposed by senior figures who challenged the legitimacy of these targets where existing priorities were already in place. The board then agreed that the EYC should support CPP-level priorities, rather than KCAs.

The Extranet was also abandoned at a very early stage. PMs were frustrated by the inability of practitioners to upload TOC data independently of PM input, or use the structure to actively communicate with one another. Consequently, TOC leaders who had accessed the Extranet failed to see the relevance to their practice – none spoken to accessed the Extranet on anything approaching a regular basis. The interface was also described by all PMs as clunky and difficult to operate. For CPP 3's PM, the Extranet was inappropriately designed, assuming that TOC learning could be communicated using numerical TOC data with little contextual factors. In her experience, communicating TOCs was always challenging given that each consisted of lengthy development stages, frequent revisiting of assumptions and a multitude of PDSA cycles which depended on context for their meaning.

Key Change Events initiated across 2015 were received more enthusiastically. Several common areas of testing were underway across CPPs – notably income maximisation for pregnant women, increasing uptake for health start vitamins, and increasing uptake of 27-30 month review – where teams had very similar ambitions. Key Change Events, focussed around shared aims with similar organisations attending, appeared to be more successful in facilitating collaborative innovation than Learning Sessions: both CPPs 2 and 3 reported being able to share learning from successful TOCs, and in developing new networks focussed around specific issues which PMs considered likely to have achieved some translation of learning.

In contrast, demand for Improvement Bootcamps remained high across CPPs. PMs felt (echoed by TOC leader interviews) that Bootcamps – in common with Learning Sessions –were often irrelevant to the workloads of practitioners, drawing too heavily on clinical examples. Consequently, practitioners often emerged sceptical and lacking understanding of how tools learned could be adapted in practice. PMs in CPPs 1 and 3 described that TOCs in their experience always needed on-going improvement support across the testing process. Consequently, all CPP case studies were observed to initiate their own QI training programmes in addition to the support provided centrally.

Level of strategic support

A common theme emerging in PM interviews was the importance of visible support from leaders within the CPP and ICS. This was a practical issue, since such figures also had control over budgets and the authority needed to create institutional space for improvement work, however it appeared just as influential as a signal to give the EYC legitimacy and encourage participation.

CPP 1 failed to attract support from the Chief Executive and Head of Children's Services, who lacked enthusiasm for the EYC's methodology and for whom the Locality Model was the CPP's priority. Participants in the leadership workstream in CPP 1 complained of an inability to attract high profile figures to workstreams as a factor which led to declining interest and its ultimate demise. CPPs 2 and 3 however featured significantly more support, with Programme Managers appointed at a more senior level, and the EYC feeding into ICS partnerships more directly. Executive support was also a factor pulling key figures from across the early years workforce into leadership positions within workstreams in CPPs 2 and 3, being cited by PMs and workstream members interviewed as key factors sustaining workstreams in CPP 2 despite their lack of productivity. In CPP 3, EYC activity took up a large portion of the ICS Strategic Partnership meeting attended by the researcher, and the PM was able to attract substantial funds in financing TOCs, and funding two Improvement Advisors posts to support her workload.

Senior leadership was also a crucial factor in the survival of EYC Home Teams. CPP 2 kept its workstreams active for two years longer than the others, while CPP 3 integrated its Home Team with ICS thematic groupings. The absence of senior support was considered by a past Workstream leader interviewed in CPP 1 to have facilitated their collapse. However, senior leadership by itself was not enough to ensure productivity of the Workstreams, since in CPP 3 and eventually CPP 2, it was an observed lack of productivity which led to declining interest and ultimate disengagement despite high levels of senior support.

Approach to TOC development

Dissatisfaction with early TOCs.

Each CPP area saw an abundance of poor quality TOCs initiated prior to PM appointment, and particularly within the first Action Period. Strategic-level interviewees unanimously considered this unproductive and damaging in the long term, since it signalled poor coordination and focus to practitioners and senior leaders alike. TOCs which emerged during this period were described uniformly as poorly focussed, lacking a coherent rationale, and exhibiting a poor grasp of improvement methodology. None of these TOCs would achieve any degree of impact, with the vast majority petering out shortly after initiation.

Supporting TOC development

A significant difference was seen in the approach taken to enlisting and supporting TOC development. Many TOC leaders had only met CPP 2's PM on a handful of occasions. PMs in CPPs 1 and 3 in contrast believed all TOCs to require regular and comparatively intensive support both in project orientation and throughout the testing process. Both PMs estimated spending over half their time in direct improvement support provision at first interview.

As Table 8 shows, CPP 2 had more than twice as many claimed active TOCs compared to the others, and in early 2015 had listed over 50 active TOCs. The more hands-off approach to TOC development combined with its Improvement Bootcamps enabled the support of a far larger cohort of TOCs, though many of these appeared to meet the same fate of those emerging during the first Action Period, fizzling out not long after practitioners returned to their own settings. This CPP subsequently reappraised its approach and consciously narrowed down its active TOC profile over the latter half of 2015, concentrating support provision on those TOCs deemed to have highest potential for impact.

Manager and practitioner roles in TOCs

In CPP 1, a large proportion of the TOCs initiated were led by practitioners and team leaders, with few service managers featuring actively within TOC projects. Many TOCs had emerged from the bottom-up with PM encouragement. In CPP 2 also, a continued focus on encouraging practitioner leadership led many of CPP 2's TOCs to be initiated with limited managerial involvement.

In CPP 3 however, the PM took a more managerial approach to TOC development, relating at first interview that 'we should not be seeing tests of change which do not link to strategic priorities', where 'priorities' reflected existing strategic directives within the CPP. CPP 1 was observed to come around to this position, by final interview seeing managerial involvement as crucial from an instrumental standpoint: without managers creating the space for improvement, TOCs were not incorporated within organisations and lacked the scope to achieve impact. The fate of many of TOCs within CPP 2 would corroborate this, with the vast majority of practitioner-led TOCs failing to arrive at significant improvements.

Summary of comparative analysis findings and implications for the EYC

The application of open coding of documents, interview transcripts and observations has revealed substantial differences across two broad domains of EYC functionality discussed here: strategic management and TOC development. These are summarised in Table 9 below.

Table 9. Similarities and differences in EYC development among case study CPPs

	Similarities	Differences
Strategic management	Collapse of collaborative infrastructure Workstreams and Extranet failed Limited engagement with KCAs and pioneers Stronger pull of CPP-level, rather than Scottish Government, priorities Continuation of the away team: Maintenance of the away team Better engagement with Key Change Events	Senior support and leadership Senior leadership absent in CPP 1 EYC seen as a lower priority in CPP 1 Pre-existing ICS infrastructure and priorities EYCs integrated with differently organised ICS infrastructure and strategic trajectories
TOC development	A more focussed approach to TOC development Dissatisfaction with TOCs emerging prior to appointment of the PM	TOC Alignment with managerial priorities CPP 3 sought managerial support before beginning testing; CPPs 1 and 2 were more supportive of practitioner-led TOCs – though CPP 1

	Conscious attempt by late 2015 to focus support a small number of high-potential TOCs	later moved towards ensuring managerial sign-off before testing.
	Each TOC sought additional QI support within CPPs to complement Scottish Government and PM efforts	Approaches to workforce development
		CPPs 1 and 3 invested in building central support, while CPP 2 invested heavily in mass QI training for its workforce (CPP 2)

The total collapse of EYC workstreams and the Extranet, and the limited success of KCAs and Pioneers across case study CPPs strongly suggests that this was not the right structure for CPPs. The demise of workstreams was the second significant failure of the EYC, after early TOC development in the first Action Periods had fallen flat in the absence of strategic initiative and dedicated technical support. These factors were closely associated with the value of the EYC being increasingly located within its methodology, rather than its function as a network or platform of collaboration.

TOCs also notably converged on the position that a more focussed and intensive support to TOCs was required, involving continuous improvement advisor support and coaching from PMs or other experienced staff. PMs were intensively involved in all of their best developed TOCs, and easing TOCs into improvement methodology was considered an intensive process. PMs in CPPs 1 and 2 estimated that perhaps 15 TOCs in a full-time case load would be possible to support, but this would have left no time to carry out other PM duties. This suggests that initiating the EYC without dedicated QI support in place within CPPs was a strategic mistake, and that the QIC model should not be considered a resource-light approach to service reform.

A final notable trend is the divergent strategic trajectories of EYC chapters despite increasing Scottish Government attempts to steward development. ICS priorities and existing infrastructure were observed to substantially alter the course of development and constrain the potential for change. A danger of this approach is that national attempts at steering can have no influence on local development. There was little evidence of KCAs or Stretch Aims influencing TOC development approaches – a key factor within the EYC's 'theory of change' laid out in the Stock Take evaluation (Scottish Government 2014). However, some success was seen in Key Change Events which focussed efforts around distinct thematic areas of priority to *both* CPPs and the Scottish Government. Here, the alignment of national and local priorities appeared more conducive to collaborative innovation than the broad-based Learning Sessions and Workstreams.

Significance of findings

For the development of the EYC

According to its interim evaluation (Scottish Government 2014), the EYC was on good course with the timescale of its anticipated development, claiming that around half of active TOCs were considering scaling up. By learning session 7, the EYC was continuing to book out its Learning Sessions, and could point to 650 reported active tests of change underway across all 32 CPPs. Data collected from CPP case studies paints a less flattering picture. If TOC development figures from Table 8 generalised, only around 20 of those 650 would be able to demonstrate any significant or scalable improvement, most only to small number of other sites. Data collected from Change Funds, from CPP story boards and websites, and in the views of case study CPP PMs of other CPP areas, would accord that this was a general experience. In sum, the

empirical data presented here suggest improvements made by the EYC remained far too limited to have feasibly had any impact on its Stretch Aims.

Empirical data from across case study CPPs highlights a number of limiting factors which restricted the scope of improvement activity, the success of this activity, and most significantly failure for learning to spread beyond its originating context. Mitigating factors prominently included:

- the failure of many TOCs during the first Action Periods,
- the collapse of Workstreams,
- the failure of the Extranet,
- the failure of Stretch Aims or Key Change areas to significantly alter the character of TOC emergence,
- the limited success of Key Change Events, inhibited by their late introduction and paucity.

Taken together, these criticisms show that the EYC failed to function effectively as a learning system. Collaborative innovation within and across CPPs was very weak, with no example discovered across the EYC of one TOC transferring across CPP boundaries, a key assumption within the EYC's design (Burns 2015), limiting the potential for change to the incremental and local rather than the transformational and national.

At present however we lack a consistent understanding of why findings have emerged: is it that failures in the implementation – the presentation of the model, the delivery of the programme, or the approach of CPPs in taking the EYC forward – which have led to underperformance, or are more fundamental structural and methodological problems to blame? The Scottish Government's merging of the EYC and RAFA into an integrated improvement collaborative, and its Stock Take Evaluation (Scottish Government 2014), makes plain that the problems are considered the former. Without an account of the operational level of the EYC however – where the EYC's

methodology is put into practice – we cannot assess the viability of the EYC's methodology. Regarding RQ3, it is not currently clear if or how the BSC model could adapt in order to be more effective, and what areas of theory and practice can learn from this. Regarding RQ1, lacking an understanding of the experience of improvement projects, we cannot address the interlinked function of learning, self-organising and coordination capacities.

Conclusions

This chapter draws from strategic-level and aggregate operational-level data to begin the empirical analysis of the EYC. The comparative analysis surfaces a number of important similarities and differences. CPPs shared notable experiences, including a widespread collapse of the home team, integration with CPP, rather than national, priorities, and a settlement that fewer, better supported and more focussed TOCs was the most effective strategy. CPPs also showed notable divergences, including the EYC's level of priority and senior leadership within CPPs, its perception of value and the approach take to TOC development. The comparative analysis strongly suggests that the EYC underperformed relative to its initial hopes, suggesting in regards to RQ3 that additional barriers may be expected in applying QICs to the population level. However, it is not possible to determine the viability of the large-scale QIC through strategic-level data alone. To shine light on this issue, the following chapter undertakes an in-depth analysis of the operational level.

Chapter 6: improvement and learning in the Early Years Collaborative: insight from the operational level

Introduction

This chapter presents a detailed analysis of operational-level data from embedded TOC case studies to generate a better understanding of the improvement process. It aims to clarify how learning and improvement have emerged in practice, and to situate the position of the EYC's structure and methodology in facilitating or constraining the improvement process. This chapter contributes a richer and more fine-grained account of the QI process outwith clinical settings than currently exists in the literature, contributing significantly to RQ 3. It complements Chapter 5's elaboration of strategic-level data, setting the stage for the application of the theoretical model in following chapters.

Introduction to TOC case study data

Research data

Findings in this chapter are derived from four data sources:

- Preliminary interviews with 4 TOC leaders and a focus group with six TOC practitioners.
- 11 in-depth TOC sub-cases situated within the three case study CPPs.
- PM perspectives on TOC development derived from six semi-structured interviews and unstructured and informal conversations.
- Secondary data, including substantial project documentation of other improvement projects (roughly 20 projects), 8 TOC leader interviews conducted as part of CPP 2's EYC evaluation, and a survey (N=22) of CPP 2's Home Team.

Each TOC case study was embedded within a distinctive institutional context, and was characterised by a unique and often complex development process. This called for generating what Ryle (1949) called a 'thick description' of each case, which explains not just behaviour, but the context on which its meaning depends (Lincoln and Guba 1985). In reporting this data, a context-rich summary of between 1000 and 1500 words was produced for each TOC, detailing the historical and contextual factors which led to its creation and influenced its development. It was not possible to report these within the thesis itself due to the combined length (around 15,000 words) however fuller summaries are available from the researcher on request.

Two summarising tables are presented. Table 10 reports an abridged summary of the eleven TOC projects, featuring key contextual characteristics, notable developments and impacts. Table 11 then reports a descriptive summary of the key TOC features which this chapter explores. Appendix G reports further descriptive statistics of the wider TOC population comprising service contexts, TOC development, and KCA alignment. The remainder of the chapter will assess the TOC development process across the main stages of the improvement model:

- The planning process: setting aims, measures and generating ideas for improvement
- The testing process: using the PDSA cycle drive service achieve improvement
- The scaling process: taking successful improvements to scale

Table 10. TOC case study summaries

TOC 1: Rainbow Programme

This TOC was a new project led by NHS Allied Health Practitioners delivering an interactive 2-hour workshop to family-facing professionals, which aimed to impart key child development lessons around nutrition, communication and motor skill development. The TOC was led by three staff members, with significant PM involvement and senior manager support. Project aims of knowledge take-up, quality ratings and satisfaction were measured through a questionnaire delivered immediately following programme delivery. Feedback and suggestions were solicited on the same questionnaire, which were discussed at team meetings and was used to continually refine programme delivery.

The MFI was applied systematically and consistently to measure effectiveness over a twelve month period. All measures showed improvement over the course of a twelve month period, though measures were erratic and could only be sensed in the long-term. Attendance was increased through accessing Locality Model networks of early years workers, which was achieved through the advocacy of the PM. By late 2015 there was evidence of user suggestions falling in number, leading staff to believe possibilities for improvement were exhausted. Instead, more fundamental changes in delivery were pursued. Connections were made with a local third sector organisation to deliver 'doorstep' lessons directly to parents. Buoyed by its positive results, by data collection termination plans were in place spearheaded by the senior manager to spread learning to other CPPs, with four CPPs in talks to extend the programme.

TOC 2: Stay and Play

The TOC was embedded within a Locality Model initiative developed in response to a community need for safe places for family activities. The project adapted an existing model of 'Stay and Play' active within a neighbouring CPP, a programme of after-nursery play activities for families using nursery playgrounds out-of-hours. Two community workers associated with the Locality Model led the project, who the PM had put through an NHS-run QI training programme. Two nursery officers were involved in project delivery, however did not engage significantly with the improvement project despite encouragement from TOC leaders. The

initial project aim was to increase and sustain session attendance, chosen because it was simple to measure and could unite the nursery and community workers around a simple goal.

The project ran over two school terms, over the course of which gradually increasing numbers of families had engaged, which it was able to demonstrate through a run chart. There was significant variation in attendance rates due to weather, periodic user access, and disruption from term times and holidays. In place of attendance data, observations of user behaviour and conversations with families had intimated that certain changes had led to improvements (improved visibility, better play activities, and improved structure of activities), and this feedback was used to initiate continuous improvements in services. Despite the somewhat positive trajectory of the TOC, the Locality Model ceased operating in late 2015, and project workers were re-assigned, leading to the cancellation of the TOC. Despite there being another Stay and Play centre in the CPP, no evidence of learning was understood to have spread.

TOC 3: What Matters to You?

This TOC was a new project led by a senior officer in Children's Services, which aimed to improve the responsiveness of services to user preferences. The TOC leader considered children's services often poor at identifying user priorities, and hypothesised that engaging parents systematically in conversations about their experience in services would generate data which could be used to re-design services at a policy (e.g. feeding into ICS delivery) or a practice (within individual services) level. The TOC leader began testing by discussing with friends and family who were parents, with conversations articulating valuable experience of service success and failure, which the practitioner reasoned could have been used to improve services. Such conversations were formalised into structured 'What Matters' conversations, consisting of four questions about their views and needs as a parent.

Arrangements were made with the Locality Model to stage two large community involvement events, which were attended by equal numbers of early years workers and parents. These generated 55 responses, which reflected two common factors: poor access to enjoyable, affordable social activities in the area, and a lack of social connectedness leading to loneliness and isolation. The project was marked by a complete absence of PDSA cycles despite its practitioner being highly trained in QI. The infrequency of interventions led to an impossibility of producing meaningful run charts, or conducting standardised observations with the MFI.

Lacking senior support within ICS however, there was little evidence that responses gathered had impacted upon service redesign. Only when the TOC leader was given the opportunity to feed into a new CPP-wide Parenting Strategy one year later was there any evidence of study aims being achieved. By time of study termination, the 'What Matters' conversations were in the process of moving into nursery settings, where

they would take place in a more focussed and systematic fashion with parents of children falling short of their 27-30 month reviews.

TOC 5: Child Development Centre

Based within an NHS-run family support centre, this TOC aimed to increase parental engagement in an existing staff-led play, development and therapeutic sessions. Staff had noted a lack of meaningful parental engagement in staff-led activities, which was considered a key problem since parents would not draw any transferrable skills from the sessions to apply at home. The MFI was adopted by the team manager TOC leader and the small staff team in a reflective manner, disregarding its call for measurability.

In total, at least 15 PDSA cycles were initiated across a period of a few months, with unstructured practitioner observation and user feedback being used to ascertain effectiveness. Throughout this process, practitioners learned that amongst other things, parents lacked an understanding of centre aims, enjoyed the social opportunity to chat with other parents, and found various aspects of the environment distracting. Action on these issues resulted in gradual impact over the following few months to achieve definite impact, though practitioners considered there was still some way to go by study termination.

TOC 6: Family-based Model

This TOC was embedded within one multi-agency locality centre delivering the CPP's 'Family-based Model', a major local authority-led approach to inform local services redesign. The centre was involved in delivering various activity, play and therapeutic sessions and engagement activities within its local area with a community development focus. The TOC's aim was to improve attendance across all of its parenting programmes, which was felt to be an aim which all partners involved could contribute to.

The TOC leader operating independently but with supportive management had initiated around ten PDSA cycles, however the model was not applied systematically and repeated measurements were only apparent in one service focussed on engaging young parents in community services. The practitioner initiated a series of interventions to test how best to send reminders and encouragement to attend. Several variations of this (numbers of texts, timing of texts) were tested, however no improvement was detected in attendance rates. However, by discussing the intervention with parents the practitioner found that being reminded had made

some parents feel valued and motivated them to attend, and that one variation, texting the day before, was universally preferred by attending parents despite not manifesting in quantitative improvements.

TOC 7: Engaging Vulnerable Families

This TOC was a new project initiated by an NHS paediatric physiotherapy practitioner. The TOC leader was invited to attend CPP 2's QI Bootcamp by her management at short notice, and hadn't fully developed her ideas for a project aim. During the process of working with an IA, the practitioner was inspired to refocus on better engaging the most vulnerable service users who were the most likely to not engage with the service. The physiotherapy service was offered on an opt-in basis, with several means of access: home visits, nursery appointments, and appointments within the physiotherapy service centre. The TOC leader knew from previous interactions with casework families that many of this hard-to-reach group were uncomfortable with strangers entering the home. The practitioner felt that making families aware of alternatives to home visits, e.g. nursery or centre-based appointments would improve uptake.

The TOC was not integrated into the wider service, with no institutional space was created for carrying out the improvement project. This was compounded by an infrequency of referrals from deprived areas, and a lack of team involvement. At time of first interview, only two observations had been collected, and no inference could be made as to effectiveness. Shortly after, the TOC stalled and ran aground, with additional burdens on practitioner time arising from staff shortages. As a result, the TOC did not progress to a higher level of development and was largely abandoned by April 2015.

TOC 8: School Literacy Project

As part of the School's strategy to close the attainment gap, the School Head had given the go ahead for a trial project extracting third year boys showing deteriorating engagement with classes from normal class to a specialised literacy-focussed programme. The TOC leader was a senior teacher and had accessed RAFA support and Bootcamp training to take forward this new project. After identifying target boys, consultation with teachers identified that target children, lacking basic literacy skills, would disengage from learning when faced with large blocks of text. The intervention was therefore designed to extract target boys from normal class to a special focussed literacy programme to help them build basic core literacy skills.

The TOC leader adopted a student-centric approach under the presumption improved involvement of target users in lesson design would lead to better engagement with learning. The TOC leader engaged students in

conversations about their reading and writing habits, finding that they did employ key literacy skills in their own settings, for instance in reading football magazines. The TOC leader then incorporated these materials within lesson plans, finding that boys did sustain engagement when source material was changed. Initial success of the project was inferred by the continued attendance of the boys at the non-compulsory class over a number of weeks. At end of first interview, plans were in placed to continue classes, and measure satisfaction and literacy skill development with standardised testing.

TOC 9: Income Maximisation for Pregnant Women

This TOC was a new multi-agency project which aimed to increase uptake of a telephone-based local authority-run money advice service for pregnant women. On project initiation there were very low referral rates to the service stemming from midwifery, despite the presumption that midwives would signpost to the service. Before project initiation, the TOC made contact with both services to arrange buy-in, and a Driver Diagram was developed linking the project to CPP-level strategic priorities around poverty reduction. Testing was led initially by one midwife who had signed up, who then began testing her ideas to increase uptake. These began with a series of service marketing efforts in which the service was presented differently to women however this was met with limited success. The practitioner began consulting with women in her caseload, some of whom suggested that having the service make first contact would improve uptake. Following this process change, immediate significant improvement was observed.

Satisfied that improvement had occurred, the project team began to work to scale the project. One other midwife was signed up from another ward, and again positive results were observed, with over 90% of midwifery referrals stemming from these two practitioners' caseloads. After some effort spent pursuing managerial buy-in, the TOC leader was able to convince midwifery managers and team leaders that the intervention was successful. Positive results enabled the TOC leader to achieve buy-in from midwifery team leaders, scaling to the two midwives' teams, and then four across all areas of the CPP by time of study termination, each time with similar improvements observed.

TOC 10: Engaging Dads

This TOC was embedded within an existing multi-agency project aimed at developing services available to male carers in the area. Led by one senior practitioner, the TOC aim was chosen as improving male engagement in nursery activities and focussed on one nursery to do this. A second area of testing was also

initiated to improve the quality and uptake of a set of 'dad's cards' aimed at delivering takeaway lessons in child development to male carers.

Initial testing involved three PDSA cycles modifying nursery service marketing to dads, but engagement rates remained vanishingly low. A consultation event with fathers staged by the service suggested that a lack of father-specific play groups had limited opportunities for engagement with services. Drawing on this learning, a dads-only play session was trialled at the nursery, which was well attended and appeared to significantly increase attendance relative to other interventions applied. Testing around 'dad's cards' attempted to measure fathers' satisfaction with the cards, however the low volume of cards being returned meant that no meaningful data was gathered, and data collection was not continued at time of interview.

TOC 11: Staying Put

This TOC was embedded within an existing intensive family-centred support service for families exhibiting sustained anti-social behaviour. The TOC aimed to streamline and improve its multi-agency service process. A process mapping exercise led by the PM had improved understanding of service processes and led to three areas being earmarked for application of the MFI. Only one area, the reduction of lengthy assessment times, had progressed to testing. Practitioner input had shown duplication and poor confidence with paperwork, with a lack of uniformity across agencies.

A standardised testing form was introduced and redrafted five times based on practitioner input. A run chart showed the number of agencies returning referral information within 10 days was steadily improving, rising from an average of two agencies to over four (of seven in total) throughout the testing period. Practitioners were in process of receiving further QI training and testing was beginning in two other areas of service efficiency at time of study termination.

Table 11. TOC case study characteristics

TOC	Project Development				Institutional characteristics			
	Existing or new project	Scottish Government rating	Improvement Observed	Status as of August 2015	Aim of TOC within service	Institutional Support for TOC	Team / Individual ownership	Leadership
1: Rainbow Project	New	4	Significant and sustainable	Spreading	Impact	High	Team	Joint Leadership
2: Stay and Play	Existing	N / A (project terminated)	Some slight	Defunct	Uptake	High	Team	Practitioner-led

3 What Matters to You	New	3	None direct	Planning	Impact	Medium	Individual	Manager-led
4 Family Pool Time	Existing	N / A (abandoned MFI)	None direct	On hiatus	Uptake	High	Team	Manager-led
5 Child Development Centre	Existing	2	Significant	Active testing	Impact	High	Team	Practitioner-led
6 Family-based Model	Existing	4.5	Some	Active testing	Uptake	High	Individual	Practitioner-led
7 Vulnerable Family Engagement	New	N / A	None direct	Abandoned MFI	Uptake	Low	Individual	Practitioner-led
8 School Literacy Project	New	N / A (RAFA)	Testing – no improvement	Unknown	Uptake	Medium	Individual	Practitioner-led

9 Income Maximisation in Pregnancy	New	4.5	Significant and sustainable	Spreading	Uptake	High	Team	Joint Leadership
10 Engaging Dads	Existing	3	Some	Active testing	Uptake	High	Individual	Manager-led
11 Staying Put	Existing	3.5	Significant	Active	Efficiency	High	Individual	Manager-led

The planning process

Practitioners using the MFI were expected to use its three improvement questions (see Chapter 3) in framing and planning projects before proceeding to testing using the PDSA cycle (Langley et al. 2009). These questions tasked TOC leaders to set aims communicating the project's ambition, to consider what changes they could make which might result in improvement, and finally to consider how such improvements could be measured and inferred from data. In lacking a 'change package' of predetermined interventions, this initial framing process was crucial in ensuring TOCs were focussed on high-value improvements. Despite this, Learning Sessions did not focus significantly on the planning or idea generation process.

The selection of improvement aims

Operational-level interviews explored how TOC projects were framed and how aims and improvement ideas were decided upon. Probing questions were used challenge TOC leaders to consider how and why certain ideas had been arrived at. Improvement aims became classified into three groups through open coding. Most notably, six TOCs sought to improve measures of service uptake, including attendance levels (TOCs 2, 6, 8, 10), or referral uptake (TOCs 7 and 9). Another (TOC 11) focussed on service efficiency, measuring assessment processing times. Three TOCs had specified an aim which sought improved impacts on service users through service quality improvement (TOCs 3, 4 and 5), while TOC 1 had made efforts to capture impact through follow-up survey. Document review of storyboards from other CPPs and TOCs presented at Learning Sessions also exhibited the dominance of measures of uptake as the focus of TOCs. Practitioners were probed in interview on why such project aims and measures were chosen, and how these squared with ambitions behind the TOC. Three areas emerged as important in framing projects: existing strategic priorities, measurability, and inclusiveness of aims.

Existing strategic priorities

An expected finding was that most TOCs (1, 2, 4, 6, 10 and 11) cited existing (organisation-level) priorities as key factors in driving ambitions. The leader of TOC 11 for instance described an existing 'model of change' on which the TOC was built around. However, this diminished the capacity for transformation: if existing priorities are insulated from the innovation process, then change remained focussed on the *how*, rather than the *why* of improvement. Nevertheless, this factor alone could not explain why particular areas of projects were chosen for improvement, or why certain measures of success were chosen over others. In the remaining five TOCs, TOC leaders also appeared to be motivated by their own understanding of what was feasible or what was important, and the desire for alteration of strategy.

Measurability

The MFI privileged a quantitative data-driven approach to improvement which hinged upon the choice of a measurable dependent variable as the project aim. The production of run charts for instance depended on multiple observations of its dependent variables over time to demonstrate a trendline or 'signal' of improvement. Nevertheless, most TOC leaders understood the MFI as applicable only to easily measurable aspects of services. As Table 11 demonstrates, many TOCs demonstrably struggled to operationalise measures of service impact.

TOCs 4 and 5 maintained an impact focus by abandoning the MFI's quantitative focus, with TOC 5's leader considering that 'we can pick something easy to measure, but that's not helping us do improvement'. Measurement difficulty also contributed to the divestment from the MFI and the EYC in TOC 4, which opted instead to rely on a traditional (qualitative and quantitative) evaluation to demonstrate impact. Even advanced projects with dedicated QI support found impact measurement challenging, with TOC 9's leader recounting, 'we've really struggled with impact data, because short of actually sitting down and speaking to a woman post birth you will not know [the financial circumstances] somebody has got'.

However, despite most TOC leaders voicing difficulty reconciling measurement and impact, all TOC case studies excepting TOCs 3, 4 and 5 did specify measurable aims. In doing so, it was clear that the requirement for measurement itself had substantially shaped the character of TOC projects to a greater or lesser degree.

Five TOCs (2, 6, 9, 10 and 11) restricted the use of the MFI to specific parts of the projects they considered suitable and easily measured, adopting other tools in measuring and improving service impact. TOC 2's leader for instance reported using established tools (e.g. the Outcomes Star) to measure well-being, but did not regard this as part of the TOC. TOCs 6, 10 and 11 also embedded the MFI in a wider programme of change in which the MFI was operationalised only in its more measurable aspects. Perhaps more troublingly, the use of probing questions at interview showed that TOCs 2, 6 and 10 considered which aspects of services were measurable *before* considering which were important to improve, meaning that measurement rather than importance was driving TOC development trajectories.

Inclusiveness of aims

TOCs 2, 6, 10 and 11 found it difficult to adopt the MFI within their large, complex and multi-agency services. This group responded in two ways. TOCs 2, 6 and 10 set an intentionally broad aim of increasing service uptake across all services delivered, since this was a goal to which diverse partners in the project could contribute. TOC 11 instead set out to deconstruct the service to understand what aspects of the project were important to improve. A process mapping exercise incorporating all agencies and practitioners instead produced a shared understanding of multi-agency roles in the complex service chain and through practitioner consultation three areas were earmarked for testing which were narrow enough to serve as discrete pathways for TOC development.

All PMs found excessive breadth a detriment to TOC quality, and in development sessions encouraged leaders to narrow their focus by considering specific areas for

improvement. CPP 3's PM incorporated a process mapping approach precisely for this issue, utilising this across all better-developed TOCs, which appeared to help clarify pressure points where TOCs could have greatest impact. The wider document review accords with this experience: it appeared difficult in every case for uptake-focussed TOC projects to infer meaningfully from quantitative data alone.

Improvement support

All TOCs cited support in setting up improvement projects as crucial in progressing to the testing stage. In TOCs 10 and 11 for instance the PM was involved well before projects progressed to testing, which was considered crucial in project framing by TOC leaders. The PM in CPP 1 was also cited as critical by TOCs 1 and 2 in helping projects plan to manage data within improvement work.

QI Bootcamps were also useful in helping to catalyse projects into action. TOCs 2, 6 and 7 had developed projects through dedicated QI training programmes, and in all cases emerged with actionable improvement projects. TOC 7 for instance worked with an IA at Bootcamp to alter project aims quite significantly, with the TOC leader reporting: *'[the improvement advisor] made it very clear that we work quite well at improving services for people that already access services, but what about the 50% of people who didn't come to your their first appointment? (…) that just kind of rang a bell with me'.* This was a particularly clear example where TOC focus was improved following improvement support, however all TOCs reported being shaped to a greater or lesser extent by PM guidance.

A universal finding was that the choice of TOC aim and the surrounding institutional commitment to improvement within a defined area was a key driving force behind the selection of interventions. This is not tautological, because a clear idea of ambition was essential in allowing the most impactful improvement ideas to emerge amongst case study TOCs. Five TOCs (1, 5, 7, 9 and 11) all cited having a clear aim as an essential precursor for generating good improvement ideas, while other TOCs with broad aims

chosen for inclusiveness struggled with improvement. In TOC 7, only when the practitioner had focussed on engaging vulnerable families did they begin to consider the particular barriers which these families might face. In the case of managerial-led TOC 11, only after a process mapping exercise had identified more specific areas for improvement were practitioner staff able to isolate improvements based on their professional experience.

The testing process: 'planning' and 'doing'

When aims and interventions were selected, TOCs were required to test interventions using the PDSA cycle. Particular emphasis was placed by the Scottish Government and PMs to the production of 'run charts' to plot interventions against a measureable project aim over time. Alongside this, workshops and PM training had emphasised the use of Driver Diagrams to construct a theory of change linking to strategic priorities (after Learning Session 4, particularly focussing on KCAs). These features would then provide practitioners with the means to both generate learning and communicate unambiguously its effectiveness in areas deemed relevant to the strategic priorities. However TOC case studies reveal that fidelity to the MFI was highly variable, as Table 12 shows.

Table 12. Adoption of PDSA cycle components

TOC	PDSA cycles	Quantitative aims measures	Run charts	Measures taken before and after intervention	Improvement observed
1	Yes	Yes	Yes	No	Significant*

2	Yes	Yes	Yes	No	Slight****
3	No	No	No	No	Marginal***
4	Yes	No	No	No	Marginal
5	Yes	No	No	No	Moderate**
6	Yes	Yes	Yes	No	Slight
7	Yes	Yes	No	No	Marginal
8	Yes	Yes	Yes	No	Slight
9	Yes	Yes	Yes	Yes	Significant
10	Yes	Yes	Yes	No	Moderate
11	Yes	Yes	Yes	No	Moderate
Totals	10	7	6	2	

****little to no impact on services

*** some improvement weakly supported by data

** improvement demonstrated by data, changes embedded

*convincing and sustainable improvements

Table 12 shows all TOCs excepting TOC 3 adopted the PDSA cycle as a reflective tool, signified by iterative PDSA cycles. Three TOCs however did not specify quantitative

measures so could not conventionally evidence improvements against chosen aims. One other TOC did not make repeated observations and could not therefore produce run charts. Of those remaining, only one had taken measurements before and after intervention in accordance with the interrupted time series design aspired to by the MFI.

Table 12 also shows that MFI fidelity and improvements achieved roughly accord, as we might expect given the primacy given to quantitative data, however TOC 5 achieved moderate improvements despite not incorporating quantitative data, and those TOCs abandoning a quantitative approach were still considered to have achieved important and actionable learning. The following section explores the improvement journey in greater depth and describes the main barriers encountered.

Lacking a solid evidence base for intervention design, the EYC hinged on the professional expertise and creativity of the early years workforce to improve the system they were embedded within through an endogenous process of continuous innovation and testing. Practitioners at the frontline were considered the main source of ideas which would lead to the improvement of Stretch Aims (Burns 2015) – indeed, even in the manager-led TOCs 1, 9 and 11, practitioner input was viewed as the best source of knowledge about service problems and potential solutions. Indeed, in TOC 11, not involving practitioners sooner in generating ideas was considered by the TOC leader to have significantly delayed progress. Probing questions in TOC leader interviews allowed further interrogation of where improvement ideas came from.

Common sense ideas

At its most mundane level, some practitioner-led TOCs, (2, 5 and 6) identified factors which could best be described as routine or common sense interventions, derived largely through a consideration of the project aim and deductions based on current practice. TOCs 2, 6 and 9, all trying to increase service uptake, initiated more

significant service marketing efforts, which all considered the most logical first step. However, these efforts were all met with at best minor improvements.

Service user contributions to the testing process

The majority of interventions which achieved significant improvements however drew more intentionally on the embedded knowledge of service users. This was observed in all TOCs except TOCs 3 and 4 – however even here, TOC 3 positioned increased user involvement as its intended aim, and TOC 4 was borne of a user-expressed desire for local affordable family leisure services. An important caveat here is that in TOC 1, intended service users were not children or families, but family-facing practitioners. Nonetheless, both groups shared a commonality in bringing the lived experience of target beneficiaries into the service design process. While user involvement was prevalent among TOCs, there was wide variation in the forms it took. These fell into three categories: active systematic involvement, active ad-hoc involvement, and passive involvement. The significance of these experiences is analysed further in Chapter 7.

Only TOCs 1 and 5 incorporated a systematic approach to involving users in the innovation process. In TOC 1, user-suggested improvements were collected at the end of programme delivery and fed in to changes made through team meetings. In TOC 5, service user insight was routinely used to inform new interventions, though through more informal methods of involvement, including gathering ideas through opportunistic conversations with users in practitioners'.

Ad-hoc forms of active user involvement was prompted in TOCs 9 and 10 by the failure of minor common sense hunches around service marketing to lead to improvement. In both cases, consultation with service users generated ideas for more fundamental service changes, which met greater success. In TOC 9, following five unsuccessful PDSA cycles testing variation in marketing strategy, the practitioner changed tactics, instead consulting women in her case load about why they thought

other women were not engaging with the service. A suggestion originating from some women of having the service initiate contact them, rather than vice-versa, served as the basis of a more fundamental service alteration which was the critical idea leading to the project's success.

Passive forms of involvement, drawing from past conversations or observations of user characteristics or behaviours were also important in the improvement process in TOCs 2, 3, 5, 7 and 8. The practitioner leading TOC 7 for instance knew through previous interactions with vulnerable families that engagement was hampered by an uneasiness with home visits. From this understanding, the practitioner was able to surmise that raising greater awareness of service delivery within the community might increase engagement. Observation served as the basis for improvement interventions in TOC 6, where young parent's use of phone technology formed the basis of testing a text-message based service marketing strategy, while in TOC 8, practitioners' observation of pupils' disengagement with learning following being faced with large blocks of text had informed the design of the literacy programme being tested within the TOC.

The testing process: 'studying' and 'acting'

The PDSA cycle encourages reflection on a continuous and iterative basis: observing and interpreting intervention effects (the 'study' phase), before building learning into the 'plan' phase in the following cycle. Meaningful outcome data is thus crucial to the 'study' and 'act' phases, since performance information about intervention effectiveness is needed to inform future actions.

Quantitative data

Both TOCs 1 and 9 used regular measurements over a long period of time to demonstrate sustained and significant improvements in their aim measures. TOC 1 for instance delivered over 20 programme sessions over a one-year period. Even here

however, improvements fell short of indicating a 'signal' of improvement (five consistent measures of improved performance over time). TOC 9 was able take advantage of pre-existing referral data making it straightforward to compare referral data from TOC and non-TOC sites to demonstrate comparative improvement. Quantitative data was also valuable in the cases of TOCs 9 and 10 in demonstrating intervention failure. In both cases, the lack of significant improvement observed over repeated testing using different variations of service marketing arrangements prompted a change of tactics – instead, this prompted a deeper search for reasons for lack of uptake through service user consultation.

Most TOCs however found it impossible to generate meaningful inference about intervention effectiveness using quantitative data alone, struggling in some way with data insufficiency. Service delivery in some service settings was often too infrequent to accomplish the rapid PDSA testing at the heart of the MFI. TOC 7 for instance had only two referrals come through from areas of deprivation in a month of testing, leading the practitioner to question the suitability of the approach for her chosen aim. Services delivering weekly programmes (TOCs 2, 6 and 10) also complained of the length of time needed to generate sufficient observations, since low service volume and high variability in attendance numbers meant little useful inference could be derived about intervention effectiveness. TOC 3 was unable to action the PDSA cycle at all since 'interventions' had only occurred on a handful of occasions throughout the year. Natural variation in uptake due to factors such as poor weather (TOC 2), or service users dipping in and out of service access also made the isolation of impact from variation problematic. Finally, the continuity of some TOCs (TOC 2, 4 and 8) was also naturally disrupted by nursery and school terms, with services in TOC 2 breaking over summer after eight sessions.

Qualitative user feedback

TOCs which struggled to infer meaningfully from quantitative indicators found qualitative data more suitable in understanding and responding to impact. User feedback also emerged as an important facilitator of this process, in two forms: observation of user behaviour (TOCs 2 and 5), and direct consultation with affected service users (TOCs 2, 4, 5, 6 and 8).

While TOC 6's run chart showed no improvement in attendance rates, consultation with users about service changes revealed that parents preferred one particular variation of its intervention (sending text reminders the previous day). Similarly in TOC 2, while quantitative data showed only marginal gains in attendance, conversations with attending parents indicated that interventions aimed at improving visibility were indeed improvements, again leading the change to be adopted across the service. User feedback complimented quantitative data in TOC 8, where children's viewpoints of changes were collected to understand satisfaction with the intervention, and in TOC 4, where deeper understandings of impact on self-confidence were sought through interviews with parents and children.

In TOC 5, qualitative staff insight was relied on entirely as a feedback mechanism. Staff observation of parents' responses to changes, and conversations with parents about changes were used to gauge impact and inform action on a continuous basis, with qualitative insight supplanting quantitative data in a formalised PDSA structure. While TOC 5's leader did agree that proxy indicators could have been used (e.g. scoring of engagement based on an observation protocol, or parent-reported measures of engagement), those methods were judged to miss the vital cues signifying improvement in engagement which staff communicated at team meetings. The initiation of a longer coffee break was for instance judged to reduce parent conversations in development sessions, which was detected through a 'sense of calm', an intangible quality which felt important but was difficult to measure. Judging changes based on these perceptions was seen as more valuable and insightful than prescribing rigid measures beforehand.

There was also some evidence that user feedback enabled reflection on the relevance of project aims. The aims of TOC 2 for instance shifted away from the attendance measures which served as the focus of its run charts, towards softer impacts upon wellbeing and quality engagement outcomes. The TOC leader described how qualitative feedback clarified the unintended impacts of the services:

'Parents have told us the children are less difficult and they're not showing as much challenging behaviour. They're willing to go to bed when they should sleeping all night - that kind of thing (…) we never really thought about what were the parents going to get out of it (…) we were just thinking that children were getting an opportunity to play but actually the parents are using it as a social exercise.'

- *Leader of TOC 2*

In TOC 2's experience, initial staff expectations of service value were misaligned with their service users, but became understood through qualitative insight. Later interventions, including the introduction of structured play activities then became focussed on improving service quality, rather than merely attendance.

Factors influencing the testing process

While there is data to support that the MFI can lead to significant and sustained improvements in early years settings, TOCs in the case study CPP population were far more likely to become defunct or abandoned the MFI than to progress to achieve significant improvements. Several important factors across the institutional characteristics of TOCs and the improvement support attracted were observed to influence this unwelcome trend.

Institutional factors

TOCs were located across diverse organisations involved in early years' services, each of which had their own organisational structures, cultural norms, and approaches to

change management. Unsurprisingly, there were different levels of cultural and institutional affinity observed regarding the MFI and EYC. TOCs exhibited a large degree of variation in two key areas in their institutional configuration: the extent of support for improvement activity within the organisation, and the relative roles of practitioners and managers in carrying this forward. TOC success depended on a supportive institutional environment in both areas, being cited in interview as a key success factor in the most advanced TOCs (TOCs 1, 5 and 9), while being partially held responsible for delays and lessened impact in TOCs 2, 3, 7 and 8.

Teams and management support of improvement projects

The majority of TOCs were embedded within a team delivery setting, with a number of practitioners a team leader or service manager embedded within a larger organisational hierarchy. In four TOCs (TOCs 6, 7, 8 and 10), practitioners were left entirely alone to implement TOCs, while in one (TOC 11), a single manager was leading the TOC after struggling to attract practitioner interest. The most successful TOCs however tended to be owned and shared by whole teams with both managers and practitioners engaged (TOCs 1, 2, 5, 9 and 10).

Two projects (TOCs 7 and 8) were carried out by individual practitioners in environments where no other staff member involved had QI training. This made it difficult for these practitioners to report the intent or progress of TOCs to other individuals within their organisations, leading in both cases to a sense of isolation. While TOCs 6 and 10 were individually led, management understood the function and purpose of TOCs, and consequently each was able to adapt certain components (Driver Diagrams and run charts produced in both cases) into organisational functionality, with both documents being used in management reporting arrangements in both cases.

In contrast, where TOCs remained separate to core organisational functioning, they faced an uphill struggle to achieve improvements. The slow rate of new referrals

coming in for TOC 7, the key difficulty in getting testing underway, could have been significantly ameliorated had the TOC incorporated other practitioners' case loads. However, the TOC leader considered that the MFI's complicated terminology put off other practitioner colleagues, who did not perceive its relevance to their practice. TOC 2 was complicated by being run across two different services, only one of which had MFI training. This group also reported difficulty in finding time to reflect with nursery workers as a group on TOC progress, which diminished any sense of joint ownership of the project.

Shared ownership was universally considered as important to project success. In TOCs 1 and 5, daily communication through team meetings provided opportunities for reflection on progress and collaboratively deciding on changes introduced. Joint involvement with the TOC also enabled a sense of shared ownership to develop which deepened enthusiasm for the both MFI and the wider project in which it was embedded. Both projects benefitted from regular contact and co-location: the walls of the office of the project team leading TOC 1 for instance were covered with run charts and whiteboards displaying planned changes. TOC 9 stated that when midwifery teams became involved, there was a palpable sense of energy released and improvement escalated sharply. In TOCs 6 and 10, the use of Driver Diagrams also helped develop a sense of shared purpose across services.

Managerial support

Another crucial institutional factor was managerial support. Managers firstly had a key role in making space for improvement projects during the testing stage, both culturally, through creating an environment of acceptance for failure and learning, and practically, by reducing practitioner delivery duties to make room for improvement work. However, managerial support was not universal. In TOC 7, the practitioner had to incorporate the TOC on top of her existing workload. This was compounded by staff shortages and intense workloads, and eventually the practitioner struggled to

contribute even an hour a week to the TOC. Not surprisingly, this TOC struggled to gain a foothold in the service and eventually ceased functioning.

Improvement Support

All TOCs accessed some form of improvement support prior to initiating TOCs. Many attended Improvement Bootcamp cohorts, and some had also accessed local QI training opportunities. All TOC leaders had a practical understanding of designing and implementing the two phases of the MFI and had familiarity with Driver Diagrams, producing run charts and often other QI tools such as process mapping. Nevertheless, most TOC leaders cited difficulty with at least some aspect of implementation of the MFI during the testing process. Consequently, most TOCs placed heavy demands on improvement support from PMs, with TOCs 1, 2, 6, 9, 10 and 11 all receiving frequent (monthly at least) contact with PMs.

PMs had a key role in presenting QI tools to practitioners in a way which was supportive, and not additional to, practitioners' work. TOCs 7 and 8, lacking dedicated PM support, reported the MFI being an extra burden. TOC 2 in contrast reported that continued PM support and some practical tools dispensed (e.g. a TOC reporting template) had supported her work: 'It didn't feel like any extra work, and it's actually given more ownership to everyone'. PMs recognised the danger of the MFI being viewed as a managerial tool or flavour of the month, and each had significantly altered their approach to TOC development over time, toning down technical language and introducing QI components to suit the specific problems practitioners were addressing. In contrast, participants at two-day Bootcamps experienced a crash-course in improvement which left some operational-level interviewees with more questions than when the first attended.

While each TOC required continual support, there was evidence of TOC support needs lessening as practical experience and expertise increased. TOC 1 for instance was competent enough to lead improvement independently by early 2015. Similarly, TOC 9

was led by an increasingly competent team with support from midwifery managers and experienced practitioners. The PM reported this growing familiarity had enabled her to be more hands off regarding the testing process with a number of TOCs, particularly TOC 9.

The scaling process

A key assumption underpinning the EYC was that learning generated through the use of the MFI could be taken to scale (Burns 2015). Scaling 'up' within the BSC model is accomplished through the iterative use of the PDSA cycle, building from a small number of users or areas to larger numbers. Scaling 'out' is accomplished through the adoption of successful interventions in other sites, communicated through Learning Sessions, or e-learning environments like the Extranet. Without this stage, improvement would remain isolated and incremental, and the EYC would have failed to function as a learning system.

However, the extent of scaling observed amongst case study TOCs remained marginal. No TOC in the case study sample had drawn learning from other projects they were exposed to, while PMs could not point to a single incidence of interventions being adopted from other CPP areas however TOC 1's manager was in conversations to scale up the programme to three other health boards. This finding appears to be replicated significantly across CPPs. By study termination, there could not be identified a single incidence of significant learning spreading across CPP boundaries.

The use of MFI in scaling improvements

Of the two case study TOCs which achieved significant improvements, only TOC 9 conformed to the scaling process envisioned within the MFI. This project began working with just one midwife, and only after demonstrable improvements were attained was the TOC extended to another midwife. After improvements were achieved again, the improvements were then spread to the two teams in which these midwives were embedded. By time of data collection termination, the project was

operational in all four localities of the CPP. In contrast, the scaling process for TOC 1 was linear, rather than geometric. There was only enough demand for the training programme to support one team, and instead the only route to scale was expanding across CPP boundaries.

The challenge of contextual uniqueness was overcome somewhat by the approach taken by CPP 1's PM in actively transposing changes across different contexts. TOC 1 for instance extended its provision by partnering with a local third sector organisation to deliver 'doorstep' sessions, taking parts of the programme and delivering them directly to parents. This partnership was facilitated directly by the PM who established connections between the organisations and played a role in project development. The PM was also in discussions to transpose TOC 3 into a nursery setting, capitalising on the opportunities created by the receipt of Attainment Challenge Funding, moving 'what matters' discussions between services and families into nursery settings where children had failed parts of the 27-30 Month Child Health Review. All of this spread however drew from the PM acting as a facilitator and network manager, and consciously advocating for the TOCs underway within the ICS partnership.

Learning from the EYC was also used to inform policy development in CPPs 1 and 3. In CPP 1, information collected about parenting needs as part of TOC 3 was being used to inform a local authority-wide parenting strategy at time of data collection termination. There was also evidence of in CPP 3 of the EYC's approach informing the development of a continuous improvement strategy shared across ICS.

Factors affecting the scaling process

Project ambition

The assumption that TOCs would aspire to share learning or adopt changes in other settings was not guaranteed. Besides TOCs 1 and 9 which had actively shared learning, only three other projects (TOCs 3, 4 and 8) had intentions to share learning with other actors. Each of these projects was initiated with the express intention of producing

learning which could inform service change beyond the scale in which it was practiced. The leader of TOC 8 considered that her intervention – a programme of classes focussing on core skill development – had potential for application in other school departments. In particular, she considered the approach to hold potential for improving numeracy skills, even before the project had generated evidence of success.

Other TOCs (5, 6, 10 and 11) did not have significant ambition to scale learning beyond their service boundaries, viewing the MFI primarily as a tool of local service improvement. These cases shared a focus on improving existing services, rather than consciously embarking on a larger programme of change. TOC 5 for instance had a clear aim of better parental engagement and could demonstrate improvement, but did not have plans to share learning with other family centres, viewing the MFI purely as an intra-organisational improvement tool. All of these TOCs were built into larger pre-existing services which carried their own service plans and budgets which were considered off-limits, no matter what learning was generated.

Variation in service delivery arrangements

Learning from TOCs 1 and 9 was spread to very similar service environments run by the same organisations (NHS settings in both situations). TOC 1 was engaged in initial conversations to spread into three other CPP areas, led by the strategic manager. The presence of similar Allied Health teams in these CPPs helped senior managers envision how the project might spread quite straightforwardly. In TOC 9, the presence of community midwifery teams across regional areas enabled a gradual scaling process across similar sites. As the project grew, community midwifery managers took notice and midwifery teams became more eager to take part.

Most TOCs (excepting TOCs 1, 2, 5, 7 and 8) took place in environments which were unique or highly non-standard across CPPs and health boards, which limited the potential for direct replication across settings. Three TOCs (3, 4 and 6) did not have clear analogues in operation across CPPs, while three others (TOCs 9, 10 and 11) were

multi-agency partnerships which were either unique (TOCs 10 and 11), or configured very differently in other CPPs (TOC 9).

There was however evidence of thematic alignment across TOCs, evident in clusters of TOCs around priority issues. TOC 9 reported that presenting its work on TOC 9 at a Key Change Event had led to continuing informal conversations with a number of CPPs with similar ambitions. Even in this instance where communication was established around a shared goal, the spread of learning took place in a highly uncertain and non-linear way. Speaking about the experience of sharing learning from TOC 9, the PM reflected:

'When you look at it from the outside it's really difficult to see whether or not you've had an influence, but actually you have because they haven't made the mistakes that you made (...) but it's very hard to put your finger on that and to kind of say that was a result of the collaborative.'

- CPP 3 PM, final Interview

The TOC leader reported sharing key pieces of learning, including recruiting midwifery team leaders to lead change, and being quicker to move on when positive results were seen, to which other CPPs appeared receptive. The experience of both TOCs 1 and 9 was that similar services can be arranged in a very diverse way across CPPs, in terms of both the services which were available, and the organisational roles involved, and thus conversations and relationships, rather than just data, were crucial in spreading learning from TOCs.

Institutional support

Since scaling projects up invariably required resources supporting the change process, the scaling process hinged on significant institutional support stemming from an increasingly large amount of actors. While some TOCs had attracted managerial support for scaling up, many had not discussed the options for scaling with managers

(TOCs 2, 5, 6, 7 and 8). Strong managerial support from the allied health lead was crucial in championing TOC 1 and leading to conversations about scale with counterparts in other health boards. In TOC 9, greater institutional resistance was encountered in marshalling support for a social intervention in a sector traditionally concerned solely with clinical responsibilities. The failure to enlist midwifery managers sooner led to delays in scaling up the process change to midwifery teams, even when the intervention was demonstrably successful.

The scaling process depended just as crucially on practitioner support. In TOC 1 all practitioners understood the MFI as an empowering tool, affording them the opportunity to take ownership over the project. In TOC 9, spreading amongst midwifery teams depended on the advocacy of one midwife who convinced sceptical colleagues that the MFI did not represent any extra work or additional duties. The spreading process then relied on resolving the concerns of both managers and practitioners: the former that improvement was feasible and reliable, the latter that it did not entail extra work and was instead an empowering tool.

Quantitative Data

The use of the MFI as a tool for proving impact using quantitative data appeared to be a pre-requisite for scaling to occur in the case study TOCs and the wider CPP sample. Quantitative data demonstrating improvement was vitally important for convincing decision makers to adopt changes in both TOCs 1 and 9. Collecting qualitative data over one year was held by the PM and the TOC leader to strengthen TOC 1's claims of improvement and its appeal for other health boards. In TOC 9, quantitative data was critical in scaling within a more difficult institutional environment:

I produced a report in September which I sent to midwifery management which basically told them that around 90% of all referrals from midwifery were coming from these 2 midwives [in the TOC] because I kept being told oh no, no, we're all doing it! No, no they're all supposed to be doing it'

TOC 9's leader was able to use this data to illustrate the benefits of an alternative approach, which overturned managerial assumptions of how the services they managed actually operated. This was a crucial turning point in the TOC's development, with midwifery teams across the CPP signing up shortly afterward. Reflecting on her experience supporting TOCs at final interview, CPP 3's PM believed the MFI was most valuable in helping practitioners document their activity, providing actionable data which better linked planning and practice. In this view the MFI's ability to 'prove' impact was central to its wider 'improving' function.

This conception of the MFI as a proving tool was common among managerial participants in the EYC (notably TOC 11) however it was also a draw for practitioners. TOC 8's leader, when asked if she had opportunities to share learning with colleagues and managers, responded: 'I'm sorry but it's not like that. I've had to sell this'. Practitioner-run TOCs 2, 6, 7 and 8 were all drawn to the MFI as an opportunity to generating data which could evidence improvement and gain the support of decision makers for service change.

Significance of findings

The empirical account of the PDSA process in this chapter has described how MFI has been used to achieve improvement in this new context across three integral stages of the improvement process (planning, testing, scaling) involved in translating individual innovation ideas into systemic improvements. While there was no universal development path which TOCs followed the MFI in moving towards impact, four prominent factors seem to be shared among the most successful TOCs, and be lacked by less successful TOCs.

- **A supportive institutional context is crucial**

All case study TOCs led by individuals struggled to gain a foothold in organisational environments. Practitioners leading TOCs suffered from feelings of isolation and lacked the resources needed for testing, while managers leading TOCs alone lacked access to the creative energy and service knowledge of practitioners. Every TOC which achieved impact did so through the combined effort of engaged practitioner teams and supportive management. TOCs had to be seen as integral to organisational practices, rather than additional to them, and there is likely to be little value in encouraging TOC emergence on a completely decentralised and ad-hoc basis amongst practitioners.

- **Qualitative data were beneficial in encouraging improvement – but quantitative data were more important for scaling improvements**

Infrequency of observations on aims measures, high natural variation and low service volumes, all combined to limit the utility of quantitative data in the PDSA cycle as a reflective tool in perhaps the majority of service contexts. Most TOCs instead opted to use qualitative insight – staff observation or direct user feedback on interventions – to gauge improvement more quickly and with greater interpretative power. However, quantitative data was more powerful as a 'proving' tool, allowing TOCs to demonstrate improvements over time, and win the support of decision makers in scaling up improvements.

- **Service users were integral to the improvement process**

Service user input was detected in some form in the majority of TOCs, and was linked to many of the most significant improvements observed. User involvement was incorporated in the PDSA cycle prospectively in the 'act' phase, in generating ideas for interventions from service users, and retrospectively in the 'study' phase in order to better understand intervention effects. While some TOCs (1, 3 and 10) adopted formal methods of involvement (consultation events and questionnaires), it was more common for informal methods, including practitioner conversations with service users

on caseloads, and observations of users behaviour to be adopted on a flexible and ad-hoc basis.

This finding is significant because service users have no formal role within the MFI or the BSC itself, nor was their involvement encouraged by the Scottish Government at Learning Sessions or otherwise. While QI is often billed as an inclusive process (Bataladen 2007), it is notable, given its alignment with the Scottish Government's public service reform agenda (Housden 2014), that service users have no formal place within the EYC. The finding is also significant to the research aim independently, since it links service users through arrangements of co-production to the innovation process and outcomes-focussed service improvement. This observation forms the basis of RQ 2, and the rationale for Chapter 7 in which this relationship is interrogated in greater depth.

- **TOCs needed improvement support on a continuous basis across the improvement process**

A final factor was that TOC success depended significantly on PM support throughout the improvement process. This was crucial in supporting TOCs to develop meaningful aims in project orientation, in providing technical QI assistance during the testing phase, and to act as a network builder and project advocate in actively sharing project learning during the scaling phase. PMs acted to overcome some common barriers and concerns, including setting overly broad aims in the testing phase (TOCs 7, 9 and 11), the perception of the MFI as extra work (TOCs 2, 10 and 11) and in actively facilitating the transfer of projects across different service settings (TOCs 1 and 3).

Contributions to RQ3

Chapter 5 has asserted that the EYC failed as a learning system, but could not determine if this was a superficial problem of implementation and presentation, or if more fundamental structural and methodological problems with the QIC model itself.

The in-depth analysis of the EYC's operational level in this chapter has shown that while the MFI can lead to improvement in an early years context, contextual conditions often limit the effectiveness of its methodology.

The need for knowledge generation led TOC leaders to rely more heavily on qualitative data, drawing often from the experience of service users themselves, to generate the most impactful service innovation ideas. Four factors stood out as particularly in expediting improvement: creating a supportive institutional environment, using qualitative data for 'improving' and quantitative data for 'proving' and scaling interventions, the involvement of service users in the PDSA cycle at both 'plan' and 'study' phases, and the need for intensive and continuous improvement support. These findings are explored further in Chapter 8, which undertakes an assessment of the QIC model's viability through an application of the theoretical framework.

Conclusions

This chapter has contributed a context-rich exploration of the improvement process within TOC case studies. In so doing, it has contributed a more fine-grained analysis of the MFI in practice in a population-level outcome context than currently exists in the QIC literature. The experience of case study CPPs across the three stages of the improvement journey (planning, testing and scaling) suggest that the EYC struggled with key methodological challenges, which alongside technical difficulties with implementation and presentation prevented any genuine learning system from emerging. Regarding RQ3, problems with measurement, variation in service context, and institutional and cultural factors which manifested suggest that the QIC model faces significant additional barriers operating within a multi-agency social service environment which have not been anticipated by QI literature.

The discovery of user involvement as a crucial factor in the testing, but not planning, stages also provides an important starting point for an analysis of how user involvement might be harnessed for outcomes-focussed system transformation, which

is explored further in Chapter 7. More generally, this chapter has also provided a more solid empirical foundation from which to begin more focussed exploration of the research questions. Drawing on the full range of data across both strategic and operational levels, Chapters 7, 8 and 9 will now begin more focussed inquiry into the thesis' research questions.

Chapter 7. Applying the Complex Systems theoretical framework: the co-production of learning and improvement

Introduction

The exploration of operational-level data in the previous chapter has found that while learning and improvement was achieved, it often occurred through substantial deviation from the process envisioned by the MFI and the BSC model. This chapter begins the application of the theoretical framework developed in chapter 3, and contributes a more fine-grained and theoretically-informed analysis of the learning process. The chapter draws on the experience of the Locality Model in CPP 1 as a comparator case (Weiss 1995) to explore the role of service user co-production in generating learning across single, double and triple-loop levels. The chapter then reflects on the significance of the findings for co-production research in fulfilment of RQ2.

Learning and improvement in the TOC case studies

Chapter 6 has introduced the eleven TOC case studies and summarised the main improvements which were achieved. This section begins a theoretically-informed exploration of this empirical data through the application of the Complex Systems framework introduced in Chapter 3. Table 13 categorises the areas of learning noted across single, double and triple-loop levels, and notes the significance of improvements which emerged as a result.

Table 13. Learning and improvements observed in TOC case studies

Learning and learning impacts	Single-loop		Double-loop:		Triple-loop:	
	Area of learning	improvement observed	Area of learning	improvement observed	Area of learning	improvement observed
1: Rainbow Project	User preferences	Service Content ***	Opportunities for collaboration	New service creation *		
2: Stay and Play	Process effectiveness	Service content and process improvements **	Unanticipated service impacts	None		
3 What Matters to You					User priorities	None◊
4 Family Pool Time			Service impact	None		
5 Child Development Centre	Process effectiveness	Modifications in service delivery **	User behaviours and needs	Creation of new processes**		

6 Family-based Model	User preferences	Process modifications*	Unanticipated service impacts	None		
7 Vulnerable Family Engagement					Service focus	None
8 School Literacy Project			User behaviours and motivations	Creation of new service**		
9 Income Maximisation in Pregnancy	User preferences	Process modifications *	User motivation and behaviours	New service process creation***		
10 Engaging Dads	User preferences	Process modifications *	Unmet user needs and behaviours	New service creation ***		
11 Staying Put			Appreciation of process failures	Process changes**		

* - no improvement

** - minor improvement

*** - major improvement

◊ - changes in progress and potential for impact at study termination

Single-loop learning

Single-loop learning was apparent across most TOCs, and this was translated in several TOCs into demonstrable service improvements. These TOCs had set goals within their areas of influence, and had chosen measurable aims such as improvements in attendance, uptake, or processing times. With well-defined outcome measures, these TOCs focussed learning efforts around existing assumptions about problems, which Greve (2003) would term a 'problemistic' search. This was manifested through common-sense adaptations, for instance the service marketing efforts undertaken in TOCs 2, 6, 9 and 10 to improve service uptake. None of these efforts achieved a significant improvement in uptake.

Better results were attained through the incorporation of single-loop learning harnessing user feedback systematically in service improvement. In TOCs 1, 5 and 9, a combination of observation and informal conversations with service users elucidated factors which led to a greater alignment of services with user preferences and achieved resulted in significant beneficial service alterations. In the case of TOC1, the number of user suggestions on questionnaires completed following programme delivery diminished over time, which signified to TOC leaders that the potential for user input in improving the design and delivery of the programme had become exhausted. This resulted in the team moving away from a focus on single-loop learning, and beginning to search for new ways of delivering the service directly to parents.

Double-loop learning

Double-loop learning was also plentiful in the TOC sample, concerned with improving understanding of the relationship between service outputs and target outcomes. In two cases (TOCs 2 and 6), feedback from service users indicated that unanticipated outcomes emerged from actions. In TOC 6, while texting families beforehand was used as a reminder to attend assuming that forgetfulness was driving poor attendance, parents fed back that this approach had made them feel more valued and brought an increased commitment to the service which was not manifested in quantitative data.

In TOCs 1, 9 and 10, the failure of initial single-loop learning efforts in refining service presentation led TOC leaders to question the assumptions behind their change theories. In the case of TOC 9, after a year of collecting and acting upon user suggestions, user suggestions decreased indicating high levels of satisfaction, and TOC leaders felt they had exhausted the potential of new user insight leading to additional improvement. This resulted in the team exploring other ways beyond the service provided to improving their goals, which led to a new service delivered directly to families through partnering with a local third sector organisation. In TOCs 9 and 10, changing how the services were marketed and presented to parents resulted in very limited improvement and helped TOC leaders understand that lack of awareness was not at the root of poor service uptake. In both cases this prompted revisiting assumptions about user motivations and behaviour, and to exploring ways forward through user engagement. TOC 9 accomplished this through conversations with women in practitioners' caseloads, while TOC 10 drew on learning collected through a large consultation day involving local fathers. In both cases, insight from service users altered the understanding of the problem, and suggestions originating from users were implemented, leading to dramatic improvements.

Double-loop learning was also put into practice in TOC 5 through normalising more open reflection with families about service effectiveness. The Initial assumptions of the team leading the TOC were that parents had poor understanding of their role within development sessions and were not aware that they were intended to

participate, however user feedback revealed this to be just one issue among many. Through staging conversational 'engagement events' aimed at exploring the issue of poor session engagement and through more informal conversations, parents revealed a conflicting desire to socialise with other parents rather than engage with sessions. Once modifications were made to eliminate this and other practical factors, the service began to experience more significant improvements in parental engagement.

TOC 8 was formed to explore how to combat increasing levels of disengagement known to be behind a worsening of the attainment gap. This involved a period of consultation with teacher colleagues who suggested that poor literacy skills lay behind disengagement with learning. Children were often observed to disengage when faced with large blocks of text, which led to a vicious cycle of stagnating literacy development and worsening engagement. This had given the practitioner a deeper understanding of the problem isolated to a particular occurrence, through which the TOC leader was able to design a targeted intervention – a literacy development class which extracted them from regular lessons.

Triple-loop learning

Triple loop learning was evident in just two TOCs, and was not translated converted into any improvements by data collection termination. In TOC 7 the TOC leader attended a locality-led QI training session on engaging vulnerable families at which they were encouraged to consider how she might focus on attracting non-engaging families in areas of high need but poor demand. When the project aim was amended, the practitioner found herself in a privileged position to draw upon previous interactions with families and generate hunches about what might be excluding them. In this instance the resonance of an external goal altered not only the strategies employed, but also the practitioner's understanding of service priorities. As Chapter 6 has discussed however, an inhibitive institutional environment prevented this re-

prioritisation from impacting at the organisational level, and eventually to the abandonment of the test.

In TOC 3, focussed user involvement around parenting experiences gathered 55 responses from parents to a structured 'what matters to you' conversation. These responses featured a high demand for affordable local family activities, and also that many parents lacked social connectedness which contributed to feelings of loneliness and isolation. The TOC leader felt that these priorities were not however shared at a senior level within ICS. However, as with TOC 7, a lack of senior support for the TOC led to little actual change. These responses were at time of study termination in the process of being fed into a CPP-wide parenting strategy.

Summary and contributions to the Research Aim

Single-loop learning was common across the TOC case study sample, which as the theoretical framework predicts resulted in alterations in performative functions or service delivery procedures. This led in TOC 1in particular to gains in efficiency and better alignment with user preferences; although in others (TOCs 6, 9 and 10) where assumptions about user characteristics or needs were faulty (at least initially), it led to very limited improvement. The application of the MFI as a purely quantitative approach in this way led exclusively to an autopoetic loop of inward-looking rationalisation.

For those that did devote energy to challenging assumptions (prompted by the failure of single-loop learning efforts to generate improvements in TOCs 1, 9 and 10), a significant amount of double-loop learning was also evident in the TOC sample. Double-loop learning was accomplished where qualitative insight was accommodated in the testing process. This issued from a re-assessment of understandings of service user behaviour or motivations (TOCs 5, 9 and 10), which led to the creation of new service approaches and significant process modifications. It also led to the

identification of unintended but beneficial service impacts (TOCs 2 and 6), and the formation of new alliances which opened the potential for new impacts (TOC 1).

Triple-loop learning was comparatively rarer in the TOC sample. Where accomplished, it resulted in a better understanding of user priorities (TOC 3), and a re-orientation of service focus (TOC 7). However, in neither of these cases did this learning challenge the priorities held at higher managerial levels, suggesting this form of learning will encounter significant challenges activating an adaptive response. This factor was only apparent in the isolated incidences where TOC leaders subjected the relevance and meaning of service priorities before proceeding to testing.

These findings make two major contributions to the Research Aim. Firstly, regarding RQ 3, the theoretically-informed analysis suggests that QICs operating in social settings need to focus on capturing and utilising qualitative data, particularly stemming from user input, to achieve the transformational change necessary for outcomes-focussed system transformation. The theoretical analysis points to a failure among the EYC leadership's of the mechanisms facilitating and constraining improvement. The continued presentation of the EYC as a quantitative data-driven approach to improvement, coupled with its emphasis on technical knowledge and QI training, while perhaps valid in clinical environments, was unsuitable to the social conditions in which TOCs operated.

Secondly, regarding RQ2, the findings show that service user feedback was an important force in driving single and double-loop learning, and was critical to the most significant improvements observed across TOCs 1, 5, 9 and 10. User experiences were used to drive improvements through single-loop learning in TOCs 1, 2 and 5; users were also harnessed for double-loop learning to alter understandings of service impacts in TOCs 2 and 6, and in actively contributing suggestions which led to dramatic improvements in TOCs 5, 9 and 10. The survey conducted as part of the evaluation of CPP 2's EYC is also illustrative here. While the survey was small (N=22) and biased

towards senior operational managers, the results strongly support that user involvement is often an implicit component of improvement work. While only 10% agreed that service users influenced their improvement work through direct forms of user involvement, 65% reported some form of 'indirect' involvement. Similar nods to the importance of user input could be detected in TOC project documentation collected from outwith the TOC case study sample.

Yet while co-production was prevalent in the testing stage, it was conspicuously absent in the planning stage. With the partial exception of TOC 3 which disengaged entirely from the PDSA cycle format, no instances of co-design or co-prioritisation were found, nor were user experiences drawn upon to improve understanding of outcome formation. This leaves a gap in understanding relative to the theoretical framework, which attests to the importance of triple-loop learning in situations where target outcomes are poorly understood or contested. This is problematic for the research aim, since we lack a coherent empirically-grounded understanding of the potential for the theoretical framework in practice. Consequently, as Chapter 4 has explained further, the research approach was extended to take account more fully of the Locality Model's (CPP 1) experience, which was founded to explore and respond to the priorities of parents and families with its locality. Interviews were conducted with integral members of the Engagement team driving the reforms initiated by the Locality Model, alongside individuals involved with its leadership, and document review was extended to take account of internal reports, progress reviews and service information provided by these individuals.

The experience of the Locality Model

The Locality Model was a major local authority-led service reform effort in children's services developed over 2011 and implemented in early 2012 . It was founded with a dual remit: to find out what priorities families living in the area had, and to test new integrated approaches to service delivery, which would be tested and rolled out across

the city. The Locality Model was championed by the Head of Children's Services and the council's Chief Executive, and was granted a ring-fenced budget over three years to explore, test, and finally roll out new models of service delivery across the wider CPP area. The Locality Model was modelled on the Total Place approach developed in English local authorities (HM Treasury 2010). Notably, no detailed strategic plan of reform was constructed *a priori* to the Locality Model's initiation; rather its ambition was for its priorities to be led by community needs and aspirations.

In early 2012, family-facing services from across the statutory sector were brought together for an intensive 'hot-house' programme of collaborative planning and community engagement, which was led by a service design agency. This also involved a large-scale survey of the local area taken previously to the event, and on-the-spot consultations taken on the locality's main high street to assess user priorities and feelings about services in the area. These insights revealed (in common with the learning generated through TOC 3) dissatisfaction with the availability and accessibility of local family services and a demand for affordable, safe opportunities for family recreation. While these were factors already acknowledged by Locality Model leaders, interviewees noted that community involvement legitimated their place as strategic priorities and provided a more solid basis for action.

The Leadership Group, staffed by ICS leaders across the public and voluntary sectors and led by the Head of Children's Services, then drew up a plan to respond to this identified need, which in line with the Total Place approach focussed on the exploitation of existing community assets. Three services would be introduced over the following two years which would address these needs:

- In June 2012, a 'Family Pool Time' service was proposed which provided low cost access to the local swimming pool. This service would prove very popular and was evaluated positively over 2014-15. TOC 4 would become embedded within this service on the EYC's initiation in January 2013.

- In 2013 a school holiday provision service was also initiated inspired by consultation with local families, which had highlighted school holidays as particularly difficult to find affordable activities. These were well attended, with around 30 families and 60 children attending sessions delivered across 2014.
- A 'Stay and Play' service (later becoming TOC 2) initiated in 2014 would address the lack of play park facilities in the area, borrowing a model in which nursery play facilities were opened for use by families after-hours.

The Locality Model was also leading other longer-term projects in the area, notably an ambition to set up a community centre in the area, and engaged in concerted efforts to develop stronger networks amongst disparate early years services in the area. However, growing internal dissatisfaction with the level of community engagement in the Locality Model's first year led to the funding and formation of the Engagement team underneath the Leadership Collaborative to take charge of its community engagement function. The Engagement team was led by a senior officer within the Locality Model, supported by 3 dedicated community engagement workers.

The Engagement team on initiation adopted a wide range of methods of community involvement, comprising door knocking, consultations and questionnaires, and more significant community engagement events (including two 'What Matters to You' (TOC 3) workshops). Towards the end of 2014, the Engagement team underwent a considerable shift in engagement approach, eschewing these standard techniques of consulting and involving families in service planning, and moving towards normalising engagement through naturalistic conversations in its service delivery functions.

This new approach was operationalised most significantly through the creation of a programme of free family play activities offered to families during school terms, each led by two early years practitioners involved in the Locality Model's networks. These 'small groups' sessions were delivered to just 6-8 parents, who were attracted by the

provision of fun, free activities which the Engagement team had previously noted a demand for, such as knitting and arts and crafts, which were designed also to teaching parents about child development. At the sessions, workers would engage families in what the senior officer at interview called 'natural discussions' about what issues were important to them as parents. Workers would then collate and feed this information back to the Locality Model and Engagement team, while parents could be immediately signposted to other relevant supportive services in the ward and throughout the city. These regular sessions allowed staff and parents to develop enduring and trusting relationships. The quality of information emerging from this approach was judged to be substantially deeper and more meaningful than that emerging from previous conventional approaches to engagement, described in retrospect 'pen and clipboard' forms of participation by the senior officer at first interview:

'What we found was [when families] meet you for the first time and you say "what's important to you, what are the issues for you?", they'll talk a lot about the physical community and environment, they'll talk about, like there are not very many play parks and things like that (…) but you need to take that time to build the relationships with people before they actually start saying what's important to them as a person and their thoughts and feelings and their individual circumstance'

- *Senior Officer, Locality Model Engagement team*

Once trust was in place, by 2nd interview with the senior officer this process had unearthed some unexpected issues:

- The scale of poor mental wellbeing amongst parents. The majority of parents engaging with the small groups service were disclosing poor or deteriorating mental health.

- The scale of social isolation amongst parents and the desire for socialisation opportunities with other parents, which was driving the observed demand for parenting groups and play activities in the area, rather than a demand for opportunities to further children's development.
- The significance of past negative experiences with public services. Many families had purely negative experiences of public service intervention (e.g. social work or police interventions), which led to withdrawal from and avoidance of public services in any form.

By integrating engagement based around individuals, rather than community members, providing attractive activities, and building trusted relationships between users and providers, small groups sessions had qualitatively altered the character of information issuing from parents. The issues of learning here were always of a sensitive nature, and parents would not divulge these unless trust was in place. The Engagement team found that so-called 'hard-to-reach' families were attending small-groups sessions which had not previously been engaging with the Locality Model.

These small group sessions however were unavoidably resource-intensive options for gaining community input, accessing the experience of just 4-8 parents at a session. This investment however was considered necessary since it became established that trust was a necessary precondition for successful engagement. However, in contrast to the 'hot-housing' event which led directly to the setting of the early agenda of the Locality Model, there was little evidence of this other learning translating to tangible service impact. While poor mental well-being emerged frequently in contact with families, however there was no evidence of this translating into any tangible service response.

This was in considerable part due to a detachment from the Leadership Team, which was already considering its exit strategy while this learning was emerging. Reporting arrangements between the Engagement team and the Leadership Group also

remained distant and infrequent, with the Senior Officer reporting at first interview, 'maybe once a quarter'. As such, there was no opportunity space for large-scale resource re-allocation, or for the creation of new initiatives to respond to this learning. The impetus behind the Locality Model was to roll out successful interventions across the rest of the city – however, by the time its funding ceased in 2015, no service even one year following completion achieved roll out, and only one service, Family Pool Time, survived following project cessation.

The co-production of learning for public service improvement

The experiences of the EYC and the Locality Model provide additional insight into how learning can lead to effective adaptive responses on a self-organising basis. Figure 8 provides a refined model of the theoretical framework presented in Chapter 3 to incorporate co-production as a creative force for public service organisational learning and service improvement. This draws from two domains of user experience which are harnessed and fed back into service design: *service* experience – an individual's perceptions of existing services as a recipient – and the wider domain of *lived* experience – an individual's perspective on their relevant conditions or states of being.

Figure 8. The co-production of learning
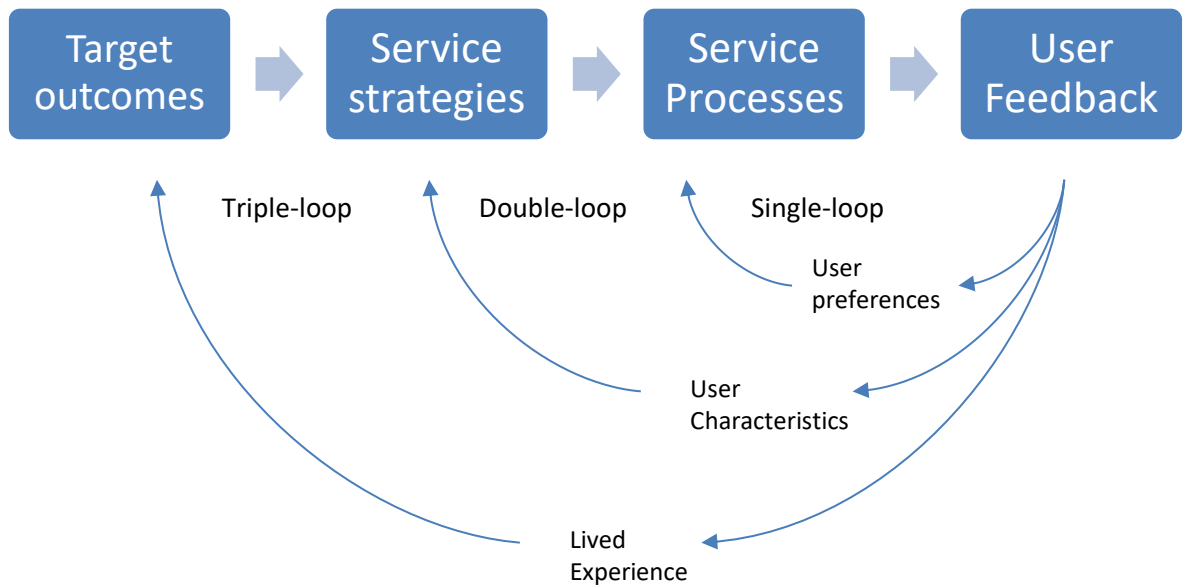
Single-loop learning across TOCs drew from service experiences, which were mediated by user preferences about routine performative functions earmarked for improvement by service providers. Single-loop learning was accomplished through feedback involving both active (e.g. consultations, discussions about changes), or passive engagement (e.g. observations of user behaviours, experience through past interactions with services). This developed feedback about interventions and suggestions for changes through which corrective changes in content (TOCs 1, 2 and 5) and processes (TOCs 5, 6, 9 and 10) enabled the optimisation of services.

Double-loop learning addressed assumptions about user characteristics – behavioural tendencies, motivations or unmet needs – which often constrained the potential for service effectiveness. Changes made were then to resolve the disjunction between user characteristics or patterns of behaviour and the design of services. Similarly to

single-loop learning, this involved *a priori* feedback about service experiences, which resulted in a better understanding of impact in TOCs 2, 5 and 6, however it combined this with an *a posteriori* consideration of unmet needs (TOCs 5, 9 and 10) which provided a more solid basis to envision new approaches and methods to implement more relevant service strategies for target service users.

In contrast to the previous two levels which focussed on the interaction of users with services, triple-loop learning concerned a phenomenological focus on lived experience of a given state of being or condition of living (van Manen 1990). Learning in the Locality Model was generated about the pressures, challenges, opportunities and priorities espoused by individuals within a particular community (i.e. parents and families within a defined geographical area). This was achieved by conscious engagement of parents and families particular which served as a springboard for system-level repurposing based on a modified understanding of outcome formation and aggregated user (as opposed to provider) priorities. Similarly to double-loop learning, both *a priori* and *a posteriori* feedback serviced this pursuit, with involvement of users in both understanding problems and visioning solutions important to the development of successful services in the Locality Model case study. However, there is evidence that feedback about lived experience can be significantly constrained through a lack of trust and the sensitivity of issues involved. This indicates that activating this form of learning entails a resource-intensive approach to building trusting user-provider relationships on the frontline.

Using the model in practice

Figure 8 provides a diagrammatic model derived from the Complex Systems theoretical framework which explicitly links forms of user feedback with learning. It postulates three forms of user feedback which link co-production with service improvement:

- Feedback about user preferences, which through single-loop learning can lead to an optimisation of service content, processes and delivery methods.

- Feedback about user characteristics, which through double-loop learning can lead to more appropriate and effective services.

- Feedback about lived experience, which through triple-loop learning can improve understanding of how the outcomes which users (rather than providers) value are achieved.

The model then provides a clear conceptual heuristic which aligns each level of co-production with its expected outcomes. In reality however it must be recognised that such choices are bounded by many confounding practical factors which constrain the opportunity space for co-production, including conflicting statutory delivery requirements, organisational hierarchies, institutional cultures, and (particularly for triple-loop learning) resource scarcity. Table 14 incorporates these considerations and provides a situational heuristic for some of the main questions relevant to public managers and decision makers designing and managing public services and social interventions.

Table 14. Harnessing co-production for service improvement: a practical heuristic

	Single	Double	Triple
When it is appropriate?	Outcomes agreed and understood, services appropriately aligned to impact upon outcomes	Outcomes agreed and understood, but uncertainty about the relevance of services to outcomes	Outcomes misaligned with user needs, causal relationships poorly understood
What can we expect to accomplish?	Better optimised services, improved efficiency	New service functions of greater impact on desired outcomes	Greater relevance of service strategy to outcomes of importance to users
How can we accomplish it?	Explore satisfaction with service functions with recipients, seek ideas on procedural modifications	Explore relevance of services to user needs, seek views on alternative solutions	Explore how outcomes are constructed from a target user's perspective (not just current service recipients)
How can we resource it?	Can be incorporated as part of service	Entails significant resource re-allocation to allow	Entails system-level-reorientation, including

	delivery efforts with little resourcing	for testing new ways of working	substantial commissioning decisions.
How can we nurture it	Practitioners must have autonomy to initiate service changes based on user feedback	Service managers must have access to significant budgets, be able to make connections with other necessary partners, and be open to trialling new services	Senior leadership must be open to the alteration of high-level strategic plans, and culturally inclined to respond to user feedback

The significance of findings for co-production scholarship

The preceding discussion has highlighted that co-production is an important and underappreciated factor within effective public service improvement in social and human-facing services. This remainder of this chapter explores the significance of this finding for the burgeoning literature on public service co-production, and considers how this extends the Complex Systems theoretical framework developed in Chapter 3. In order to situate the findings within co-production research, it is necessary first to understand the genesis of co-production and the nature of the evidence gap which a focus on learning and adaptation seeks to bridge.

The significance of co-production – the joint involvement of citizens and providers in the production of services – was first noted by Elinor Ostrom and colleagues at the

Workshop at Indiana University (Ostrom et al. 1978; Parks et al. 1981), and was extended significantly by subsequent research (for example, Sharp, 1980; Whitaker, 1980; Brudney and England, 1983; Percy 1984). Prior to this, it was understood that service users could influence the design and delivery of public services only though forms of voice or 'participation' in services (Arnstein, 1969), or through exit and choice between providers (Hirschmann, 1973).

After a fallow period during the heyday of New Public Management reforms, co-production underwent a renaissance period in public administration scholarship in the new millennium (Joshi and Moore, 2004; Brandsen and Pestoff, 2006; Bovaird, 2007), with a number of recent special editions in leading public administration journals (Bovaird et al. 2017; Osborne et al. 2016; Steen et al. 2016), and books published on the subject (Alford 2009; DuRose and Richardson 2015; Pestoff et al. 2013; Fugini et al. 2016). In contrast to when the term was first coined, co-production has also gained expediency as a policy response to modern governance challenges (Christie Commission 2011; Loeffler et al. 2013; OECD 2011; Trade Union Congress 2013).

Reviews of the state of research have noted that while significant conceptual advances have been made in understanding who can take part in co-production and in which activities, significant definitional and conceptual ambiguities remain (DuRose et al. 2017; Verschuere et al 2012). Much recent scholarship has understood co-production as an umbrella concept, and approached conceptual development through a taxonomic disaggregation across various dimensions, including who is involved in what activities or processes (e.g. Bovaird et al. 2011). This avenue of inquiry been challenged more recently by those drawing from services dominant theory (Osborne and Strokosch 2013), in which co-production is seen as an integral and inalienable component of the service delivery process, whether designed-in or not by service managers. Still others have argued that co-production should be understood normatively as a new and distinct model of public service in which power and decision

making is shared between user and provider (Cahn and Grey 2012; Needham and Carr 2008).

The lack of consensus on conceptual and definitional issues is also recognised to have fuelled a persistent evidence gap concerning the empirical effectiveness of co-production, which has been exacerbated by the lack of theoretically-consistent comparative research (DuRose et al. 2017; Needham and Carr 2009; Loeffler and Bovaird 2016; Verschure et al. 2012; Voorberg et al. 2014). Beyond this, we still lack theoretical clarity over how the pathways through which specific types of co-production might be expected to lead to specific benefits (Loeffler and Bovaird 2016; Voorberg et al. 2014).

Within recent scholarship, there has been an increasing focus on using this creative potential as an engine of innovation, improvement or transformation in public services (Bovaird et al. 2017; Bovaird and Loeffler 2016; Simmons 2016; Osborne and Radnor 2016), rather than merely a substitution for public sector inputs (Parks et al. 1981).

Work drawing on a services marketing perspective has viewed the experiential knowledge of service users as an important source of innovation (Osborne et al. 2015; Osborne and Strokosch 2013; Radnor et al. 2014). While user feedback has long been understood as driving continuous improvement (Minelli and Ruffini 2017), Osborne and Stokosch (2013) explain how service co-production can be combined with intentional design to foster an 'enhanced' co-production which can achieve more significant innovation. Needham (2009) and Needham and Carr (2008) argue from a social care context that the impact of co-production can be moved from a 'descriptive' explanation of service delivery to a 'transformational' relationship that realises service transformation through power-sharing and joint decision making.

While arguing from different disciplinary perspectives, both Needham (2009) and Osborne and Strokosch (2013) implicitly recognise that realising improvements through co-production involves a learning process for public service organisations,

which derives from the experiences and creative energy of service users and citizens. It is therefore notable that while 'innovation' and 'transformation' have featured prominently within co-production scholarship, an explicit focus on these learning processes has so far been neglected. In this paper, organisational learning theory is therefore explored to understand the relationship between learning and co-production and the implications this holds for service improvement. In doing so, the paper's focus is centred on the potential for the co-production of learning, rather than the engagement of citizens in the physical delivery of services.

Contribution to co-production scholarship

While much previous research which has hypothesised, but not demonstrated, the creative potential of co-production, this thesis uncovers very clear examples of impact at the micro-level:

- Single-loop learning was observed to facilitate a steady increase in participant knowledge uptake and satisfaction in TOC 1.
- Double-loop learning directly enabled the creation of new forms of service delivery which resulted in 90% of pregnant women referrals to a money advice telephone service coming from this method in the local authority in TOC 9.
- Triple-loop learning led to a clear refocusing of services and to the creation of a number of well-received, well-attended services under the Locality Model.

With the possible exception of the Locality Model, none of these examples would have been recognised as 'co-production' by those who were engaged with it. The instances and impacts of co-production were often hidden – either unconscious, taken for granted, or not publicised – and would emerge only when practitioners were prompted in interview to consider the sources of learning which led to improvements.

The application of the learning framework in an evaluative capacity enables a systematic and forensic analysis of co-production which makes this unacknowledged

and tacit contribution to service improvement explicit – every significant example of service improvement was resultant to user feedback. The findings accord strikingly with Bovaird and Loeffler's (2016, p.160) contention that co-production 'is a social innovation which is more hidden than publicized (…) because much co-production practice tends to emerge from the front-line, rather than top-down by organizational leaders'.

The learning uncovered therefore both verifies the utility of the adapted Complex Systems framework and extends it, allowing causal linkages to be drawn more precisely between the processes and impact of co-production, a factor called for by recent scholarship (Loeffler and Bovaird 2016; Voorberg et al. 2014). The Complex Systems theoretical framework thus provides more solid footing from which to advance inquiry into the creative potential of co-production as a key driver of public service improvement and innovation. Echoing the services management literature (e.g. Osbourne and Radnor 2015), this perspective suggests the experience of service users and communities of identity (i.e. those with lived experience of particular target outcomes) have a central role within service improvement and innovation, however it extends the focal point of innovation from an internal focus on service experience, towards an external focus on the lived experience of outcomes.

While co-production scholarship has theorised the creative potential of co-production as a driving force for service improvement and transformation, an explicit focus on processes of organisational learning has not so far entered this debate. Focussing on learning through the application of the Complex Systems theoretical framework, brings two specific benefits to this discussion. Firstly it links process and outcome more clearly than previously accomplished in the literature, providing public managers and architects of service interventions with greater clarity of why, where and how to utilise co-production for service improvement. Secondly, it extends inquiry into how the transformative potential of co-production theorised in the literature (Cahn and Grey 2012; Needham and Carr 2009; Osborne and Strokosch 2013), but not yet

significantly demonstrated in practice, might be achieved. Taken together, these contributions provide a sounder basis for public management research and practice to deepen its analysis into the creative potential of co-production.

Model limitations and suggestions for future research

This chapter has attempted to improve understanding of the potential impacts of co-production on public service outcomes. However, while it contends that a focus on organisational learning clarifies this area considerably, a learning perspective by itself can only be a partial explanation of co-production's potential impact. This chapter has shown that learning is a necessary but not sufficient condition for endogenous service improvement. Many throughputs in this process – existing strategic priorities, resource constraints, cultural barriers, and poor connections have conspired to assure that the sum total of improvement realised was less than the potential brought through the learning generated.

While learning processes and resultant innovation have been linked in organisational theory most often through double-loop learning (Easterby-Smith et al. 2000; Fabricius and Cundill 2014; Jaaron and Backhouse 2016), this chapter stops short of discussing the relationship between public service innovation (broadly, the *enactment* of novel ways of working) and public service learning, a precursor of such innovation. Finding ways of countering the 'defensive reasoning' (Argyris and Schön 1978) which inhibits organisational learning, and ways of instituting a positive error culture (Gigerenzer 2015) are equally vital to capitalising on learning if the potential of co-production to contribute to better public services is to be realised. Further research should focus not just on how learning can be enhanced, but how learning can be translated more effectively into improvements and harnessing the innovatory potential of learning.

The learning framework by itself also does not specifically address how co-production of learning can be put into practice, which remains an researched area worthy of further exploration. In practice, there are many tools which might fit this purpose: QI methodologies like the MFI may be hold potential if optimised to take explicit account of co-production; however the two literatures have remained largely separate. Service design tools such as experience-based co-design (Bate and Robert 2007) also have potential to advance co-production of learning, but work remains to be done integrating such practical methods with a structure of learning and adaptation within public service organisations.

Conclusions

This chapter responds to a growing call from research and policy for greater clarity for how the creative potential of co-production can be harnessed for public service improvement. It develops a novel framework for understanding how learning can be generated through co-production, drawing on the Complex Systems theoretical framework developed in Chapter 3. This shows how service user feedback can improve the efficiency (inputs-outputs), effectiveness (outputs-outcomes) and relevance (understanding outcomes) of public services. Taken together, the model provides a framework whereby the transformational potential of co-production, frequently hypothesised but rarely observed, can be operationalised.

Drawing on two illustrative case studies in childrens' services, the study provides rare empirical evidence tying processes of co-production to discrete outcomes. This is used to refine and extend the theoretical framework, demonstrating its potential in making often unrecognised instances of co-production explicit and showing the potential for more clearly tying process and outcomes in co-production. For co-production scholarship, the study provides a more solid basis for fine-tuned and detailed inquiry into the creative impact of co-production, including future scholarship into public service innovation, user-led improvement, co-creation and service re-design. For

public managers and service designers, it clarifies the expected outcomes of engaging with different forms of co-production, providing a practical means to understand when and where different forms of co-production are opportune.

For the Research Aim, it verifies the evaluative potential of the outcomes framework for service improvement in understanding processes of learning. The analysis suggests that co-production is a key source of learning generation in complex systems, which the Complex Systems theoretical framework suggests can facilitate the achievement of outcomes-focussed service transformation.

Chapter 8. Applying the Complex Systems framework: can QICs improve population level outcomes?

Introduction

Chapter 3 has developed a Complex Systems theoretical framework for managing for outcomes, which has been applied to operational-level data in Chapter 7. This chapter has two functions. Firstly, it road-tests the theoretical model in an evaluative capacity, drawing on the full concourse of data, to assess the potential for QIC models to be effective in population outcome improvement, beyond their role in clinical healthcare systems (Inkelas and MacPherson 2015; Scottish Government, 2012). Secondly, it aims to advance theoretical and empirical understanding of the utility of a complexity-informed approach to the improvement of public service outcomes to provide a firmer foothold for future empirical and theoretical research in public administration and service improvement. This Chapter fulfils the Research Objective 3: *to apply the theoretical model to an appropriate case*, and also provides the summative contribution to RQs 1 and 3.

Applying the Complex Systems theoretical framework

The theoretical model for outcomes improvement provides a novel framework for understanding improvement within public service interventions and programmes of reform more generally. As an evaluative model, it places focus on the enhancing adaptive capacity of systems (Lowe et al. 2016) which are achieved through the linked channels of increasing self-organising potential, strengthening performance attractors, and increasing propensity to generate and transmit learning. Features of learning and self-organisation are therefore crucial conceptual features defining the class of service interventions to which this model can be applied, and which also constitute the foci of evaluative endeavour.

Chapter 3 has argued that QICs, loosely defined here as networks of organisations which are linked by an exploratory focus on innovation towards common outcomes, provide a promising approach to the improvement of complex public service outcomes which is attracting increasing academic interest. These are argued in Chapter 3 to be promising fora for the theoretical framework both to test its evaluative potential (satisfying RQ1), and to explain the viability of these initiatives as population-level outcomes-focussed interventions. The operational methodology of the EYC – its MFI methodology and BSC structure – has been described in detail in Chapters 3 and 4, while its experience is reported across Chapters 5 and . The following section applies the theoretical model to consider the potential of the QIC model as an example of an outcomes-focussed improvement community. The application of the theoretical framework focuses evaluative attention on three capacities: learning, coordinating and self-organisation, which are discussed in the following section.

Capacity for Learning

Effective operation in complex systems requires a capacity to understand improvements in conditions of uncertainty (Snowden and Boone 2007), which is a key element of adaptive capacity (Lowe et al. 2016). The theoretical model links this capacity for learning to three areas of the service production process: process through single-loop learning, impact through double-loop learning and understanding through triple-loop learning.

The PDSA cycle is an explorative approach is not uncommon within QI literature. White et al. (2014, p.1635) for instance argue that QI approaches should aim to 'engage and empower ward teams to actively participate, innovate and lead quality improvement at the front line'. The iterative and responsive nature is highlighted by Inkelas and Bowie (2014) who argue this requires a cultural shift towards viewing data as a tool for reflection and improvement. The implementation of the MFI was observed to promote both single and double-loop learning, which led to the creation

of effective new services and significantly altered service processes across the three case study CPPs.

The lack of triple-loop learning

The absence of triple-loop learning, where learning concerns the understanding of outcomes themselves, was by a tendency to neglect a planning stage prior to initiating testing in TOC case studies, also a recognised trend in the QI literature (Dixon-Woods 2010). CPP 3's PM at first interview described the tendency for PDSAs to 'jump into testing with both feet' – a trend noted to encourage single-loop learning across TOC case studies where PM or IA input at the project planning stage was absent.

As Chapter 7 has noted, the absence of triple-loop learning is problematic since it deprives a system of its ability to challenge and refine conceptions of the complex problem systems from which outcomes emerge. As Gilstrap (2005) notes, goal-setting itself is itself a deterministic action which presumes the current direction is the correct one. For this reason, Reed and Card (2016, p.148) argue, 'an important role of the wider methodological approach is to conduct investigations prior to starting the use of PDSA to ensure that the problem is correctly understood and framed'. The complexity of factors influencing children's developmental outcomes makes it all the more necessary to initiate a reflective process on 'why' aims and changes are constructed, as opposed to merely 'how' they are to be implemented.

The effectiveness of the MFI in generating single- and double-loop learning

Chapter 6 has described that the MFI was operationalised to generate knowledge in two ways. Firstly, the MFI was used as a 'proving' model, where a focus on generating quantitative evidence as proof of change led to single-loop learning and resulted efficiencies in service processes. Secondly, it was adopted as an 'improving' model, where a reflective focus on improving services based on qualitative insight (mainly based on user feedback) led to double-loop learning and resulted in the creation of new strategies and significant process changes.

As Chapter 6 has described in greater depth, TOCs struggled to make sense of impact through a purely quantitative approach and required to incorporate qualitative feedback to make sense of impact. However, the object of a learning system is to enable learning to inform practice beyond merely the site of its origin. In order to achieve this, learning had to be reified in such a way that it would reduce the uncertainties of others in a position to facilitate the scaling of the intervention. Managers involved in TOCs in contrast to practitioners tended to see the MFI's value as a 'proving' model to improve the quality of decisions made. Quantitative data was more immediately salient in resolving uncertainty for managers and convincing them to create the institutional space for TOCs to expand. No TOC within the case study sample or the wider CPP population was observed to achieve any degree of scale lacking quantitative evidence.

While some TOCs did manage to use the MFI for both 'proving' and 'improving', in line with Miller (2014) the two logics were observed more often to run into conflict. The engagement of practitioners was repelled by the perception of the MFI as a managerial data-driven agenda, while managerial engagement was repelled by a lack of tangible evidence generated through practitioner reflection. This creates a tension between the accomplishment and communication of learning which can undermine learning capacity.

Finally, the need for measurability in improvement aims guided TOC leaders to frame projects around more easily-measured indicators of process and uptake as the focus of improvement projects, rather than impact. The theoretical model explains that this is likely to lead to an inward-focussed autopoetic transformation cycle. Insisting on the same standards of measurability therefore is likely to limit the dissipative potential of learning, relegating improvements made to service process-oriented adaptation.

Capacity for Coordination

As Lowe (2013) notes, actors cannot reasonably be held to account for outcomes over which they have little to no control. Outcomes are argued to carry greater potential for meaningful coordination as *performance attractors*, which orient actors towards outcomes but do not hold them to account for quantitative impact on outcome indicators. The coordinative capacity of a system initiating an intervention is determined by its ability to attract the self-organising behaviour of autonomous actors towards system outcomes and stimulate new, more relevant forms of organisation.

At the EYC launch event, the Scottish Government's Clinical Director's related in his plenary speech about SPSP Stretch Aims, 'I don't know if we're going to get to [the target], and the big secret is: I don't care. I care only that it changed the system'. With only very informal accountability linking CPPs to the pursuit of Stretch Aims, the articulation of the EYC's target outcomes was very closely aligned with a performance attractor function. Later KCAs were designed to function also without lines of accountability, designed to be 'impossible to resist' (Scottish Parliament 2014), and to attract the voluntary support of CPPs.

In line with BSC guidance, Stretch Aims were designed as measurable system-wide measures which could only be achieved through system transformation, rather than merely incremental improvement (IHI 2003). Some detail was given as to what the manner of this transformation would look like. Firstly, Driver Diagrams were produced for each Stretch Aim (found in Appendix H) Secondly, KCAs introduced at Learning Session 4 represented the 'big ticket items' (Scottish Government 2014b) which the Early Years Quality Improvement Unit believed would have the most significant bearing on Stretch Aims. Documentation and discussion focussed on KCAs at Learning Sessions and Key Change Events further clarified the assumptions underpinning KCAs and their drivers (Children and Families Analysis 2014).

However, the data does not support that KCAs and Stretch Aims were influential within TOC development or EYC strategic management within CPPs. Existing CPP-level priorities operating through strong vertical accountabilities functioned as a strong point attractor which Chapter 5 has shown worked to diminish the strategic importance of Workstreams in CPPs 1 and 3. The pull of CPP priorities was also observed to lessen the impact of KCAs on TOC development, particularly in CPP 3 where they were rejected in favour of existing CPP-level priorities at an ICS board meeting attended by the researcher. The lack of influence at the strategic level meant that Stretch Aims and KCAs were not communicated to the operational-level Home Team and TOC leaders. No TOC leader interviewed felt they were prompted significantly to take account of KCAs or Stretch Aims in TOC development sessions with the PM.

An additional factor however was the perceived vagueness of both KCAs and Stretch Aims, which also became apparent through interviews at both strategic and operational-level. Stretch Aim Driver Diagrams (Appendix H) gave no clear representation of where individual agencies might fit in within 'primary' or 'secondary' drivers of Stretch Aims, and which can be critiqued from a systems perspective for ignoring the fundamentally recursive nature of complex problems. As Bryk et al. (2011, p.17) note, Driver Diagrams are intended to draw 'attention to the specific hypotheses undergirding improvement solutions'. However the lack of a change package acknowledged that such solutions were absent – consequently, even secondary drivers were did not clearly communicate the position of key agencies of relevance to the Stretch Aims.

The data suggest QICs operating to achieve population outcomes will face greater challenges in coordinating activity towards shared goals owing to a stronger pull from competing strategic priorities. The SPSP was initiated and funded to a significant extent by the institutional body responsible for its delivery, had senior leadership and

sign-off from the hospitals with which it engaged, and participating teams were afforded the time and resources to carry out testing alongside existing duties. The TOC development process documented in Chapter 6 describes a different experience: many TOCs had little senior leadership input, perhaps the majority of TOCs did not receive any additional funding to support changes, and practitioners were often required to carry out testing with no reduction of existing delivery duties. The discordance between EYC priorities and organisational-level priorities led to TOCs being somewhat dislocated from core organisational operations.

Capacity for self-organisation

The autonomy afforded to agents in CAS theory enables them to undergo an adaptive self-organising response to learning which is critical in mobilising knowledge across the system. A dissipative response in this context manifests not just at within the behaviour of one agent, but as a wider systemic response in the surrounding system.

The EYC aimed to mobilise learning through its collaborative infrastructure, including the Extranet, the Learning Sessions, and most prominently, the Workstreams. However, the extent of self-organising behaviour which occurred through such platforms was marginal. Owing in large part to this, there were very few TOCs which spread to other sites within CPPs, and not one example could be found by study termination of a TOCs spreading from one CPP to another, a key assumption with the EYC's formulation (Burns 2015). The thematic analysis highlighted three factors emerged as significant in reducing the EYC's capacity for self-organisation: the failure of collaborative infrastructure, the fragmentation, autonomy and heterogeneity of context among children's services, and the context-dependency of learning generated.

Failure of collaborative infrastructure

Observations of Workstreams in CPP 2 found little evidence of any genuine collaborative innovation. While Workstream members were observed to give advice or

share experiences in meetings observed, there were no incidences discovered of collaborative TOCs emerging or achieving scale through the Workstream platforms.

The creation of Workstreams had brought together diverse agencies with different cultural approaches to change and little previous interaction, which felt arbitrary to participants and were demonstrably unproductive as observations in CPP 2 revealed. The age-basis of Stretch Aims was not viewed as a conducive rallying point for innovation by strategic-level interviewees, since Workstream members often had little understanding of one another's practice. PMs reported a similar issue with the Extranet, which was extremely broad in its presentation of improvement work and made locating relevant learning to the problems or contexts faced by practitioners extremely difficult.

A better record of knowledge transfer took place where there were clear shared goals and similar service contexts, as with the case of spread observed in TOCs 1 and 9, and feedback from attendees at Key Change Events. This suggests that what Englebart (2003) terms C-level learning on an inter-organisational basis occurs best around those problems which share what Bryk et al. (2010, p.6) call a 'family resemblance', which KCAs better represented. Where the dissimilarity of context and aims in Workstreams led them to function as a loosely connected network, Key Change Events then functioned as more of a Community of Practice (Wenger 1998). The EYC's experience strongly suggests that QICs should configure thematic and problem-based collaborative structures to harness collaboration and innovation. While such factors were the focus of one-off Key Change Events, the EYC could have capitalised more intensively on the variation of practice around common TOC themes.

Heterogeneity and contextual uniqueness

Chapter 6 has documented many structural barriers to knowledge transfer in social systems. Four TOCs (1, 3, 4 and 11) took place in entirely unique services, while several

others (TOCs 2, 5, 6) had just one or a few other settings across the CPP. Thus, even if significant and easily communicable improvement was generated, it lacked a natural route to scale. This meant many TOCs would require a much broader programme of change management to achieve spread, which the EYC was not structurally equipped to deliver.

In addition, the heterogenous way in which similar settings were delivered across CPPs was a key barrier in straightforwardly implementing solutions from other CPPs. The PM in CPP 3 reflected at first interview that, 'I suspect the thing that will get in the way is the fact that we are set up to work in 32 different ways across local authorities and 14 different ways across health boards'. This inherent variation was exacerbated by the different ways in how ICS partnerships were set up and prioritised, diminishing opportunities for straightforward knowledge transfer across similar contexts.

Stickiness of knowledge

The failure of learning to spread straightforwardly across boundaries was also due to the context-dependent nature of knowledge, which von Hippel (2005) terms the 'stickiness' of information. The experience of TOC 9 in sharing learning with other CPPs in an informal learning sharing network reinforces this point. Discussing this experience the TOC leader related:

But actually it's about the ins and outs, it's about the everyday, it's about the softer stuff that you don't want to put into...or sometimes the more controversial stuff that you don't want to put into a report but which you'd quite happily talk to a colleague about

- *Leader of TOC 9*

The above quote highlights that knowledge was much more complex to communicate than the BSC model assumes. Sharing learning involved the uncomfortable nature of being open with peers, the communication of the nuance and context-dependency of

learning, and the importance of communicating failures as well as successes as part of the sharing process. The PM at final interview gave examples of transferring key lessons – for example engaging midwifery managers sooner to facilitate spread – however when probed on this issue responded, 'it's hard to put your finger on it and say "that's the result of the collaborative"'.

The reality of transfer within the EYC was better understood as Kitson (2009, p.218) argues as an 'organic, incremental and chaotic' process, rather than the sequential and ordered form of knowledge transfer typified by the BSC model's programme of Learning Sessions and Action Periods. The BSC model's design can be criticised for incorporating what might be termed an instrumentalist view of evidence use which involves direct and clear-cut adoption of knowledge (Nutley et al. 2013), manifested in its case by quantitative run charts and project summaries presented as storyboards at Learning Sessions.

Summary: the contributions of the Complex Systems theoretical framework in an evaluative capacity

As Chapter 3 has argued, the EYC intended to achieve its Stretch Aims not through a linear or rationalist procedural approach, but through a decentralised and adaptive process of dissipative self-organisation where current practice is continually transformed through decentralised innovation and communication. Regarding RQ1, the application of the Complex Systems theoretical framework exhibits interpretive value in drawing attention to the key mechanisms and contexts which have vitiated or enabled adaptive behaviour as the QIC model interacts with a multi-agency social context.

The viability of the QIC as an outcomes-focussed intervention

Chapter 3 has noted the emergence of collaborative improvement initiatives operating to achieve population-level outcomes, which combine a focus on innovation through

the adoption of improvement methodology (Langley et al. 2009; Shojania and Grimshaw 2005) with a network structure to facilitate knowledge transfer. These have taken several names, including notably Networked Improvement Communities in education (Bryk et al. 2010), Quality Improvement Collaboratives in healthcare (Kilo 1998), and Collaborative improvement and Innovation Networks in population health (Ghandour et al. 2017; McPherson et al. 2015).

However, the efficacy of these initiatives within a multi-agency service environment as a an outcomes-focussed intervention has not been subjected to significant empirical analysis, despite advocacy in policy and recent scholarship (Chief Medical Officer for Scotland 2012; McPherson et al. 2015). As Chapter 3 has noted and others have observed (Cairney 2016), the pursuit of high-level outcomes requires QICs to operate in service areas characterised by a weaker and more ambiguous evidence base, and fewer off-the-shelf interventions to adopt. It therefore demands a focus on the generation and communication of knowledge, in addition to the traditional concern with implementation. RQ3 focuses attention on how QICs can respond to this new role. In so doing, it aims to contribute to a growing body of literature focussed on realising the benefits of collaborative innovation for the improvement of population-level outcomes (Inkelas and Bowie 2014; Bryk et al. 2011; Ghandour et al. 2017; Inkelas and McPherson 2015; McPherson et al. 2015). The application of the theoretical model in the previous section has identified three issues which QICs need to address to fit this new role.

A less measurable social world

The MFI is positioned as a quantitative data-driven tool which is modelled on an interrupted time series observational design. This hinges on its ability to represent goals by distinct quantitative indicators, to make rapid observations on these indicators over time, and to achieve a sufficient sample size to distinguish movement in the dependent variable from naturally occurring variation in services. These factors

enable improvements to be detected not from inference alone, but from a perceived more objective and rigorous statistical process. Results can then be demonstrated through the production of 'run charts' which clearly communicate intervention effectiveness.

In healthcare settings, many routine operations are studiously measured or at least are potentially measurable. However, social services invariably operate in environments where data is not systematically collected on routine operations, with service quality instead assessed through reflection and feedback. While quantitative data often plays a complementary role in such situations, qualitative data – for instance case reports in social work (Ames 1999) – provide the primary means of reflection, communication and case management. Quantitative indicators are known to perform particularly poorly in assessing impact which is intangible and unpredictable (Boyne and Law 2005; Lowe 2013).

In addition, many children's services are delivered on a weekly or even more infrequent basis, or with small numbers of participants. While a midwife in a postnatal maternity ward might see many mothers over the course of a day and could systematically document impressions, this is not possible for one community development worker interviewed, delivering attachment-focused play sessions for six parents twice a week in different community settings. The data suggest that social service systems present measurement difficulties so substantial as to push the MFI to breaking point.

Service fragmentation, heterogeneity and contextual uniqueness

The lack of spread observed across the CPP case studies also strongly suggests that QICs operating in population outcome settings face an innately more difficult task of transferring learning through self-organisation. QICs aspire to create a learning system, where partners can learn from one another's actions, not just their own. This

is achieved through the creation of some means of communicative infrastructure, in the case of the BSC model through the Workstreams, Extranet and Learning Sessions.

In clinical healthcare settings, variation in context is recognised to mitigate against direct transfer of knowledge and one-size-fits-all solutions (IHI 2003; Kilo 1998). The MFI is designed to provide a context-sensitive approach to knowledge transfer where interventions can be tested before committing to the significant resource costs of widespread implementation (Langley et al. 2009). Nevertheless, QICs in healthcare environments operate across *broadly* similar contexts: operational functions are likely to be structured similarly even though practice and the particularities of context may vary. While maternity units across Scotland's hospitals are indeed characterised by significantly different processes, priorities and approaches to practice, midwives share an understanding of their key service purpose: to ensure positive pregnancy experiences for all mothers. In addition, there are enough similar midwifery service contexts for interventions to have significant scope for spread.

In social systems however, services are likely to be more fragmented, culturally and operationally dissimilar and often contextually unique. The character of supportive services provided to looked-after young people for instance varies dramatically across CPP areas, characterised by different and even competing understandings of how to approach the staggeringly poor outcomes experienced by this group (Bywaters et al. 2014). The majority of TOCs in the CPP population located within community-focussed services were developed in response to their local contexts and had no clear parallels in other areas of Scotland. This creates the likelihood for far higher contextual dissonance relative to healthcare settings and resultantly, knowledge which is which is stickier and more resistant to transfer (von Hippel 1994). The stickiness of knowledge makes it more unlikely for like-to-like knowledge transfer to be a sufficient strategy for spread in multi-agency social systems. This leaves QICs in with fewer natural routes to scale, diminishing the potential for spread in the absence of a wider programme of change management.

Goal multiplicity and autonomy

In hospitals and primary healthcare settings, QICs are often imposed by the same institutions in which they are carried out. In QICs operating under the SPSP for instance, actors were bound by formal accountabilities to outcomes agreed by both the initiating system of the SPSP and the institutional setting of the hospital within which the programme was embedded. In QICs operating in patient safety and primary care improvement, aims relate most commonly to existing process goals likely already to be the focus of strategic management – hospital have a duty to reduce hospital-associated infections and patient mortality for instance (Royal Cornwall Hospitals 2015).

In multi-agency social systems, aims and strategic priorities differ drastically across service organisations, and accountabilities are often informal and multiple. The fragmentation of public service delivery brought about by NPM reforms in the UK has created a highly decentralised context where networks have become an increasingly important feature of the governance and delivery of public services (Laegreid and Christensen 2013). In children's services, while GIRFEC and the Early Years Framework provide the legislative basis for shared action, Integrated Children's Services partnerships have significant autonomy over how key areas of service delivery are structured, and resultantly thematic priorities for action vary significantly across Scotland's 32 CPPs.

Joining this increased autonomy is a heightened level of goal ambiguity. Goals are multiple in children's services for instance, because of the need to report jointly to the Scottish Government against GIRFEC (2008), to inspections carried out by the Care Inspectorate, and to CPPs for supporting delivery against SOAs. Children's Service Plans produced by ICS boards often make reference to all of these priorities. The presence of significant institutional autonomy of key actors including Nurseries and Primary Schools also means that similar contexts may have widely different

organisational practices and improvement priorities. QICs operating in a multi-agency context face the additional task of negotiating the construction of a shared vision across a far more contested landscape. Such a shared vision must have currency across traditional boundaries, while being sufficiently detailed to situate the position of a variety of agencies and individuals within it (Bryk et al. 2011).

Towards a population-outcomes focussed QIC

The interest in QICs as population-outcome focussed interventions has given rise to a growing body of scholarship, most of which has advocated the model (Inkelas and Bowie 2014; Bryk et al. 2011; Ghandour et al. 2017; Inkelas and McPherson 2015; McPherson et al. 2015). However, no detailed empirical work has been carried out of an explicitly-outcomes focussed QIC, nor has this body of literature engaged in sufficient depth with the conceptual differences concerning the improvement of institutional quality and population-level outcomes.

This thesis contributes a critical but constructive perspective to this scholarship. Drawing from the application of the CAS framework for outcomes improvement to the EYC's experience, the previous section has described three additional challenges which must be overcome for QICs to function effectively as population outcome improvement systems. The following section draws on the Complex Systems theoretical framework to consider how QICs might improve their capacities of learning, self-organisation and coordination in a population outcome context.

Improving learning capacity

Given the pragmatic focus of the PDSA cycle on generating 'good enough' evidence (Langley et al. 2009), there is a strong case for requirements of methodological rigour to be further relaxed when applied to social settings, particularly in the early stages of testing, to allow more rapid learning and a greater responsiveness to user needs and preferences. A qualitative focus responds better to the 'inherent messiness' (Ogrinc and Shojania 2014, p.265) of the PDSA cycle which as Tomolo et al. (2009, p.217) argue involves 'false starts, miss firings, plateaus, regroupings, backsliding, feedback, and overlapping scenarios within the process'.

Moving forward, there should be a clarification of the EYC's value proposition to its different stakeholders to make them aware of this trade-off. Managers should relax barriers for what constitutes evidence of effectiveness, while practitioners should be afforded increased autonomy within their roles and control over the services they provide. At the same time, the functioning of a learning system requires that learning is communicable, and so practitioners must incorporate an outward focus in considering how learning can be documented as a condition of participation. A realist view of improvement would maintain that improvements are 'sold' as much as they are shared – the generation of quantitative evidence may then be appropriate as a 'proving' strategy where institutional culture is more hierarchical, despite its limitations.

A lesson may be drawn from the 'pipeline' approach to evidence standards taken by Nesta (Puttick and Ludlow 2013) which in an attempt to resolve the barrier between innovation and evidence, relaxes the evidence burden relative to a project's stage of development. If and when learning curves begin to stabilise, TOCs might be coached to move towards carefully designed quantitative research with an explicit view to supporting its transferability. This approach is illustrated in the adapted PDSA 'ramp' in Figure 9.

Figure 9. An 'evidence pipeline' approach to Test of Change development
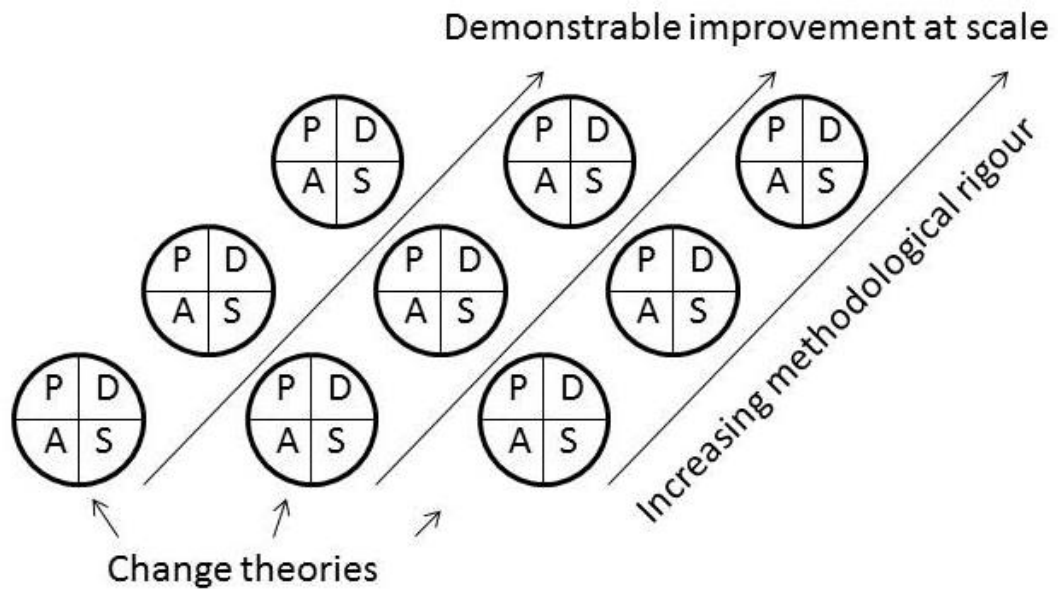
Make co-production explicit

Chapter 7 shows that co-production was among the most significant sources of learning and improvement in TOC case studies. Yet while the MFI was presented at Learning Sessions and Bootcamps to practitioners as a tool for autonomous local innovation, documentation did not explicitly acknowledge from where this may emerge. The use of service user feedback for impact-focussed improvement has been noted in QI literature (Bataladen 2007). It has also been suggested that users can act as 'quality detectives' (Bate and Robert 2007). However, while user engagement is occasionally noted as good practice in QI efforts, it remains tacit and invisible within the MFI and PDSA cycle as advocated by the IHI (see, e.g., Langley et al. 2009).

Indeed, in healthcare settings a patient's potential contribution to service improvement is likely far more limited: surgical mortality rates or the transmission of hospital-associated infections are processes where the patient's engagement is primarily physiological rather than cognitive. The intrinsic nature of co-production to the improvement process in social and relational settings however positions the differential needs, service preferences and lived experiences of populations as integral creative inputs.

Case study TOCs suggest that co-production will play a useful role in improvement work within social systems whether this are designed-in or not. However, its impact on emergent innovations may be enhanced if this process is made conscious and recognised as an essential input alongside traditional professional knowledge in an outcomes-focussed improvement process. This could be facilitated by modifying the PDSA cycle such that practitioners are prompted to collect and use feedback at both 'plan' and 'study' phases.

Reflect on outcomes in project framing

Despite the presence of three framing questions within the MFI, there was a need in TOC case studies for increased emphasis on planning and framing the project to prompt more double and triple-loop learning and thus more dissipative system transformation. Examples of improvement support at the planning stage relevant to outcomes could involve a consideration of the deterministic aspects of cause and effect relationship surrounding outcomes which are understood, a review of existing evidence, or as with the case of the Locality Model in Chapter 7, exploring how issues are framed and understood by service users themselves, which can be particularly important given the often different priorities of service providers and users (Simmons 2016). In the latter approach, some form of needs analysis might improve the coherence of service focus and user priorities. There is also the potential to integrate the systematic focus brought about by QI with the explorative and user-centred

approach of design thinking, with tools such as experience-based co-design (Bate and Robert 2007) used to set the service agenda before beginning focussed testing.

Improving coordinating capacity

A second major problem likely to be faced by QICs is the significant fragmentation of service landscapes responsible for population outcomes, which necessitate a coordinated effort among often poorly-aligned and multiply-accountable service agencies. This is recognised by Inkelas and McPherson (2015, p.1) who argue, 'transforming health care into a population health system requires methods for innovation and improvement that can work across professions and sectors'.

The data suggest that QICs should capitalise on pre-existing shared goals and opportunities for alliances which lie across systems, and ensure that performance attractors are sufficiently detailed to have resonance and relevance to the individuals at all levels of key delivery agencies. There is also a call for revisiting problem definitions and system conceptualisations on an ongoing basis to respond to emerging knowledge and ensure effective strategies are promoted.

Build a more inclusive shared vision

For Pedder and MacBeath (2008), a shared vision is essential for holding networks to common purpose, while for Senge (1990, p.344), 'building a shared vision is crucial early on as it fosters a long-term orientation and an imperative for learning'. Enhancing the potency of performance attractors can be accomplished by building a better supported vision of a desirable future system state.

The potency of performance attractors is not merely in providing strategic direction, but in encouraging agents to construct their own localised representation of a shared vision. For Gilstrap (1995) shared goals are imbued with individual meaning, giving wriggle room for conflict, disagreement and debate which can in turn change the meaning and function of the attractor (Stacey 1996). A promising alternative approach

for the QIC model is to actively co-create such a vision with integral partners in a participative visioning stage before moving to decentralised testing. A similar approach was applied in the Magnolia Community Initiative, a population outcome-focussed QIC in children's services which preceded the EYC (Inkelas and Bowie 2014), in which a system map clearly articulated the position of each contributor organisation in relation to its system-level outcomes.

Respond to motivations and regularities

The strong potency of existing priorities as a point attractor is an unavoidable aspect of operating across traditional institutional boundaries. One way the EYC could have counteracted this would have be to develop a better understanding of the strategic direction of the service system with which it sought to engage.

While KCAs did not significantly influence patterns of self-organisation, better coordinative impacts were achieved in CPPs 2 and 3 through Key Change Events. These were effective from the perspective of CPP 3's PM in motivating collaboration around thematic goals shared by service agencies, CPPs and the Scottish Government. Key Change Events provided a platform for collaboration in response to some problems which TOCs had clustered around: how to increase uptake of the 27-30 developmental month review, how to maximise incomes among pregnant women and new mothers, and how Healthy Start vitamin uptake could be increased. These areas had more potential as performance attractors since they capitalised on concise problems with existing institutional support, and created the potential for mutual gain.

Also instructive in this instance is the case of TOC 7, in which the practitioner re-oriented her practice to pursue a new priority highly resonant with her values: increasing service uptake in deprived communities. The TOC leader was able, drawing on previous interactions with non-engaging families, to generate ideas for how her service could be re-designed. It cannot be determined if this aim reflected a service

problem shared by other similar services, or if it could similarly motivate other practitioners. It may however represent a leverage point (Meadows 1999), which if exploited by the Scottish Government – through providing a bigger spotlight at Learning Sessions and Improvement Bootcamps or through its positioning as a thematic goal similar to KCAs – might initiate more significant dissipative change across service systems.

The EYC however lacked a mechanism to recognise potential leverage points and respond with the modification of performance attractors. Klijn (2008) argues that complexity requires a form of governance which is much more involved and familiar with its component parts in order to beneficially influence system behaviour. By paying attention to the factors shaping self-organising behaviour, learning might have been generated about how to promote beneficial self-organising responses.

Improving self-organising capacity

Finally, the construction of an effective learning system is more challenging in social service systems since they are characterised by higher levels of heterogeneity and fragmentation. This creates more significant context differentials which must be traversed to communicate learning. This is compounded by 'stickier' knowledge generated, and an ineffectiveness of run charts, storyboards and PDSA cycles to communicate learning or evidence effectiveness.

Foster self-organisation around shared goals

The data suggest that learning can be communicated through organic processes of self-organisation by reconfiguring collaborative platforms to better facilitate interactions. Though Workstreams succeeded initially in bringing together diverse agencies with little previous interaction, their diversity and poor focus resulted in meetings which were demonstrably unproductive. This experience recalls Wenger's (1998, p.182) warning that, 'the likelihood of irrelevance makes engagement at the

boundaries a potential waste of time and effort', as participants fail to recognise the competence or view one another as potentially productive partners.

Better collaboration took place where there were clear shared goals and similar service contexts, which facilitated transfer TOC 9 and was an enabling condition for discussions held about TOC 1's spread to other CPP areas. This suggests that the inter-organisational learning which Englebart (2003) terms 'C-level learning' occurred best around problems characterised by a family resemblance (Bryk et al. 2011). Such problem 'families' were observed to manifest as regularities of TOCs spread across CPPs. While the Key Change Events did bring organisations together for focussed collaboration around thematic priorities, these were one-off events which were not well-advertised to Home Teams.

Mobilise sticky knowledge through peer-based approaches

The EYC's cross-CPP collaborative infrastructure was very impersonal – the Extranet and storyboards at Learning Sessions encouraged the display of run charts and PDSA cycles, but could not communicate the nuance of the improvement journey which characterised case study TOCs. Since learning was embodied in errors as well as achievements, knowledge transfer was not facilitated well through the impersonal collaborative infrastructure of storyboards, run charts and Learning Session presentations. What data there is of successful knowledge transfer spread strongly accords with the view that knowledge translation is a 'contact sport' (National Collaborating Centre for Determinants of Health, p.9).

Lowe et al. (2016) argue that Wenger's (1998) Communities of Practice can encourage horizontal peer-based accountability leading to the creation of a positive error culture (Gigerenzer 2015) and better collaborative innovation in conditions of uncertainty. Dopson and Fitzgerald (2005, p.188) argue similarly that knowledge which is 'sticky at professional boundaries (…) may yet diffuse within different communities of practice'. In common with the BSC model, Communities of Practice create problem-focussed

learning environments which aim to create a shared sense of identity and purpose (Wenger 1998).

Communities of Practice bear a strong resemblance to the QIC – they share a problem-focus, a shared learning method and a distinctive learning structure. They differ however in two important ways. Firstly, they view learning as a situated, social and participative process, rather than one which is technical and instrumental. Secondly, they encourage learning transfer through a continuous joint exploration of problems, rather one which is segmented and time-limited.

Encourage knowledge transfer across boundaries

Outcomes-focussed QICs are also challenged by fragmentation and heterogeneity in service context, with few straightforward routes to scale. The experience in CPP 1 in transferring the locus of two TOCs from community services to nurseries shows that active facilitation can promote learning transfer across significant contextual disparities. Neither of these TOCs could have spread through an organic process of self-organisation: TOC leaders and nursery managers would not organically have come into contact, and TOCs themselves were highly contextually embedded and would likely have lacked obvious relevance.

In both cases, the PM operated from a privileged position as a boundary spanner (Ryan and O'Malley 2016), able to understand the potential for spread and collaborative innovation, and to respond opportunistically to new opportunities within their environment, created in the case of CPP 1 by Attainment Challenge funding. This supports Bovaird's (2008, p.324) argument that effective strategic management in complex systems requires 'swimming in the flow of events' and being alert to emergent opportunities for influence.

The data support Bryk et al.'s (2011, p.5) argument that in a population outcomes context that QICs 'need design which explicitly aims to function in the hands of diverse

individuals working in highly varied circumstances'. The EYC may have achieved a better record of transfer by placing more priority, particularly within the PM position, on active knowledge mobilisation across diverse contexts.

Concluding thoughts: what added value from a Complex Systems perspective?

The application of the Complex Systems theoretical framework has a produced a meaningful account of system performance which would go unappreciated if a purely rationalist process and outcome framework was applied. The comparatively limited progress of the EYC when compared to the SPSP for instance can be explained in two ways. Firstly, the failure of the EYC to achieve significant scale and spread might be explained through widespread deviation from a tried-and-tested improvement method, resulting from poor QI training or cultural opposition from actors unfamiliar with a scientific approach to improvement. Indeed the Scottish Government's mid-term evaluation of the EYC takes a viewpoint very close to this, with widespread QI competence as the desired end point of its theory of change (Scottish Government 2014).

The application of the framework outlines a different explanation: a failure of the EYC's management to adapt to the challenges innate to the process of improvement within a multi-agency social system. In this view, the deviation from the MFI does not arise through methodological naivety or a lack of technical capacity, but as agents struggling to resolve the innate tensions between innovation and the production of evidence.

Through the Complex Systems theoretical framework, we view key events and patterns of behaviour not just within their own contexts, but as wider systemic factors whose effects propagate across system levels and across time. For instance, understanding existing priorities as point attractors not only provides clarity over the coordinative challenge faced by the EYC as an initiating system, but allows us to better understand and contextualise the patterns of self-organisation which emerge. Thus we

understand that the decomposition of collaborative infrastructure was not just as a loss of function, but a factor both resultant from and contributing to an increasing orientation towards CPP-level priorities and away from national priorities. The limited influence of events such as the introduction of the KCAs can in turn be understood not because KCAs were poorly conceived, but because they were not compatible with these increasingly entrenched trajectories.

In sum, a complexity-informed perspective makes key processes visible which might otherwise go unnoticed or misinterpreted. The framework thus allows more meaningful explanation of the dynamic causative relationship between events and behaviours which, observed over time, explains more fully the mechanisms governing system performance.

Contributions to RQ3: the potential of the QIC as an outcome-focussed intervention

Referencing the success of QI approaches in patient safety, the 2012 Chief Medical Officer report argues 'the use of such techniques to improve population health has the potential to be equally revolutionary' (Chief Medical Officer for Scotland 2012, p.2). However, the theoretically-informed analysis suggests that QICs cannot straightforwardly replicate this success for two broad reasons.

Firstly, improvement cannot occur in social systems through the quantitative-driven technical process used in healthcare systems. The technical focus of the MFI, narrowly focussed on quantitative data and the production of run charts – was only appropriate in a relatively small subset of service contexts across children's services. Importantly, a quantitative focus could not effectively determine impact, and subsequently service processes, not outcomes, become the focus of improvement. QI methods derive from industrial production and, arguably somewhat naively (Pollitt 1996), import some assumptions from that setting: that there exists a single best way of doing things, and that this can be straightforwardly transferred across contexts. Instead, the experience

of learning and its transfer were achieved through participation and ad-hoc collaboration, and were thus better understood as a social process (Wenger 1998).

Secondly, the task of improving outcomes is fundamentally different in social systems since evidence about interventions is often contested, ambiguous, and outcomes themselves can be poorly defined (Cairney 2016). Tackling outcomes then requires a focus on the generation of knowledge, rather than just its implementation, and the deviation from current understandings of best practice, rather than just conformity to it. The Scottish Government have taken the view that deviation from the EYC's methodology limited progress (Scottish Government 2014). The value of the Complex Systems theoretical framework in this context is to strengthen claims about the actual mechanisms of improvement and thus isolate fault within the method, rather than just its implementation.

This thesis aims to contribute a critical yet constructive perspective to the emerging scholarship surrounding QICs in social systems tackling population-level outcomes improvement (Inkelas and Bowie 2014; Bryk et al. 2011; Ghandour et al. 2017; Green et al. 2012; Inkelas and McPherson 2015; McPherson et al. 2015). The analysis suggests QICs must respond with fundamental alterations in their methods and approaches which acknowledge and respond to the challenges of a multi-agency social system. Table 15 highlights three contextual challenges which featured prominently in the empirical analysis, and lists possible responses to improve the functioning of QICs in this new environment.

Table 15. Summary of QIC challenges and possible solutions

Capacity	QIC challenge	Problems observed	Possible Solutions
Learning	The social world is innately less measurable and impact, particularly about outcomes, is difficult to make sense of through quantitative data alone	A quantitative approach was impractical in many cases, and encouraged a single-loop learning process	Permit a 'pipeline' approach to TOC development, allowing more rapid and non-linear development during early stages

Systematise co-production at the 'plan' and 'study' stage during testing to enhance innovative potential |
| | | A lack of reflection before testing lessened triple-loop learning and promoted autopoetic self-organisation | Insist on planning stage before testing in explicit reference to performance attractors

Draw from other QI approaches or design thinking to systematise a prior 'TOC planning' stage |

Coordination	More significant autonomy among participants and the presence of multiple intersecting and contradicting goals	There is a greater difficulty holding together a shared vision amongst an enhanced diversity of organisations	Identify and respond to intrinsic motivations to foster collaboration around shared problems
		There is a difficulty in reconciling the high-level goals of networks with the local issues which have meaning for practitioners	Build a more inclusive shared vision through a deliberative co-creation of a shared vision
Self-organisation	Knowledge is 'stickier' at boundaries	Context-sensitivity and knowledge 'stickiness' (von Hippel 1994) made learning difficult to transfer through collaborative platforms	Foster peer-led collaboration around shared goals through Communities of Practice (Wenger 1998)
	Fragmentation, contextual uniqueness and heterogeneity of service	There are fewer natural routes to scale within or across CPPs, knowledge must travel greater contextual distances	Enhance integrative adaptivity (Bryk et al. 2011) by actively mobilising knowledge across different contexts

	delivery contexts		

Conclusion

This chapter has completed the application of the Complex Systems theoretical framework to the case of the EYC, fulfilling RO3. In so doing, it has found the framework to hold value as an evaluative tool, illuminating the function of mechanisms governing system behaviour. The chapter has contributed a theoretically-engaged analysis of the functioning of QICs seeking to improve population-level outcomes in multi-agency social service systems, isolating three additional barriers to the capacities of learning, self-organisation and coordination which QICs must overcome to be viable. Finally, the chapter has attempted to elucidate solutions to each of these difficulties, which can inform the future development of outcomes-focussed QICs and complexity-friendly approaches to outcomes management more generally.

Chapter 9. Contributions, limitations and suggestions for future research

Introduction

In this concluding chapter, the thesis' contributions to theory and practice are discussed, and consideration is directed toward how future research might build upon these. Firstly, following a reflection on study aims, the principal contributions of the thesis are outlined and discussed in the context of their theoretical and methodological limitations, and are located within current debates in relevant scholarship. This serves as a starting point for a final reflective discussion on the research aim, drawing from relevant policy and management literatures, to consider how the research and practice of a CAS approach to public service improvement can be advanced. The thesis concludes with a reflection on the implications of a CAS theory of outcomes improvement for public service management in theory and practice.

Research Aim and Logic

The thesis introduction has argued that an outcomes-focus is one of key factors driving the reform of public services internationally and within the UK in particular, and posed a dilemm for public management scholarship: how can public service outcomes can be sustained or improved in the face of worsening demand and stagnating levels of input?

Chapter 1 set out an over-arching Research Aim to structure the thesis around this problem: *to enhance the ability of public governance systems and public service interventions to achieve better outcomes in complex public service systems by advancing theoretically and empirically a Complex Systems Approach to outcomes management*. Four Research Objectives were specified to guide inquiry, and three Research Questions were developed to tackle concise gaps in relevant strands of the

public management literature which advanced inquiry into the Research Aim. The thesis logic is outlined in Table 1 in Chapter 1, while the research gaps are pinpointed in Table 5 in Chapter 4. The thesis is argued to make five valuable contributions to the research aim.

Contributions to an outcomes-focussed public management

Conceptual contribution

Firstly, on a conceptual level, the literature review in Chapter 2 has synthesised discussion about outcomes from a wide body of scholarship, including public administration, social policy, evaluation, education theory, organisational theory, public health and social epidemiology, to reveal outcomes as a contested and multi-functional concept with significant implications for the design and management of public services.

The review traces an outcomes-focus back to John Sinclair's Statistical Account of Scotland in 1791 which aimed to ascertain 'the quantum of happiness enjoyed by [a country's] inhabitants' (Sinclair 1798, p.xiii). While being accused of being an 'uncontested discourse' (Smyth and Dow 1998, p.291), outcomes can be seen to function in a number of distinct capacities within public management. Drawing from public management, social epidemiology and public health literatures, the review then constructs two broad paradigmatic approaches to the management of outcomes. Firstly, the Rationalist Approach, allied with NPM reforms, views outcomes as the 'results' of linear service production chains and combines a philosophical and economic rationalism as its core assumptions. Challenging this view is the Complex Systems Approach, which views outcomes as the emergent product of complex systems, emphasising the externality and inherent complexity of outcomes which are instead explained through complex processes of self-organisation and emergence.

This conceptual development is argued to constitute an original contribution to the public management literature, since it draws competing understandings of outcomes which have until this point been highly divergent. This puts into stark relief the level of attention afforded to each within public management theory and practice. The Rationalist Approach has been put into practice in commissioning approaches like Social Impact Bonds (Disley et al. 2011), Payment by Results (HM Govt 2011; Morse 2015) and other forms of outcomes-based commissioning (Paley and Slasberg 2007), benchmarking frameworks popularised under such as Best Value and later Public Service Agreements, or management strategies such as Managing by Objectives (Drucker 1954).

The Complex Systems approach, while being associated with systems-based design tools like causal-loop modelling, or whole-system and place-based approaches to service reform, lacks the theoretically-coherent link to practical tools which the Rationalist Approach enjoys. This makes explicit the warrant for further research within public management in developing actionable alternative approaches within the Complex Systems paradigm to begin to surmount the systematically documented failings of the Rationalist Approach.

Theoretical contribution

Building on work from social epidemiology (Jayasinghe 2011), population health (Burns 2015), collaborative service improvement (Bryk et al. 2011; Inkelas and McPherson 2015) and emerging public management scholarship (Lowe and Wilson 2016; Lowe et al. 2016), the thesis probes CAS theory to develop a constructive theoretical framework for outcomes-based management. This framework draws three fundamental components of Complex Adaptive Systems together – self-organisation, attractor states and distributed agentic learning – to outline an integrated theoretical framework which describes how outcomes might be improved through a process of dynamic system transformation. This framework links agent and system through a

dynamic process of feedback to continuously improve the fitness of the system

relative to an evolving understanding of outcomes. The framework developed in

Chapter 3 is presented again for convenience in Figure 10 below.

Figure 10. The Complex Systems theoretical framework
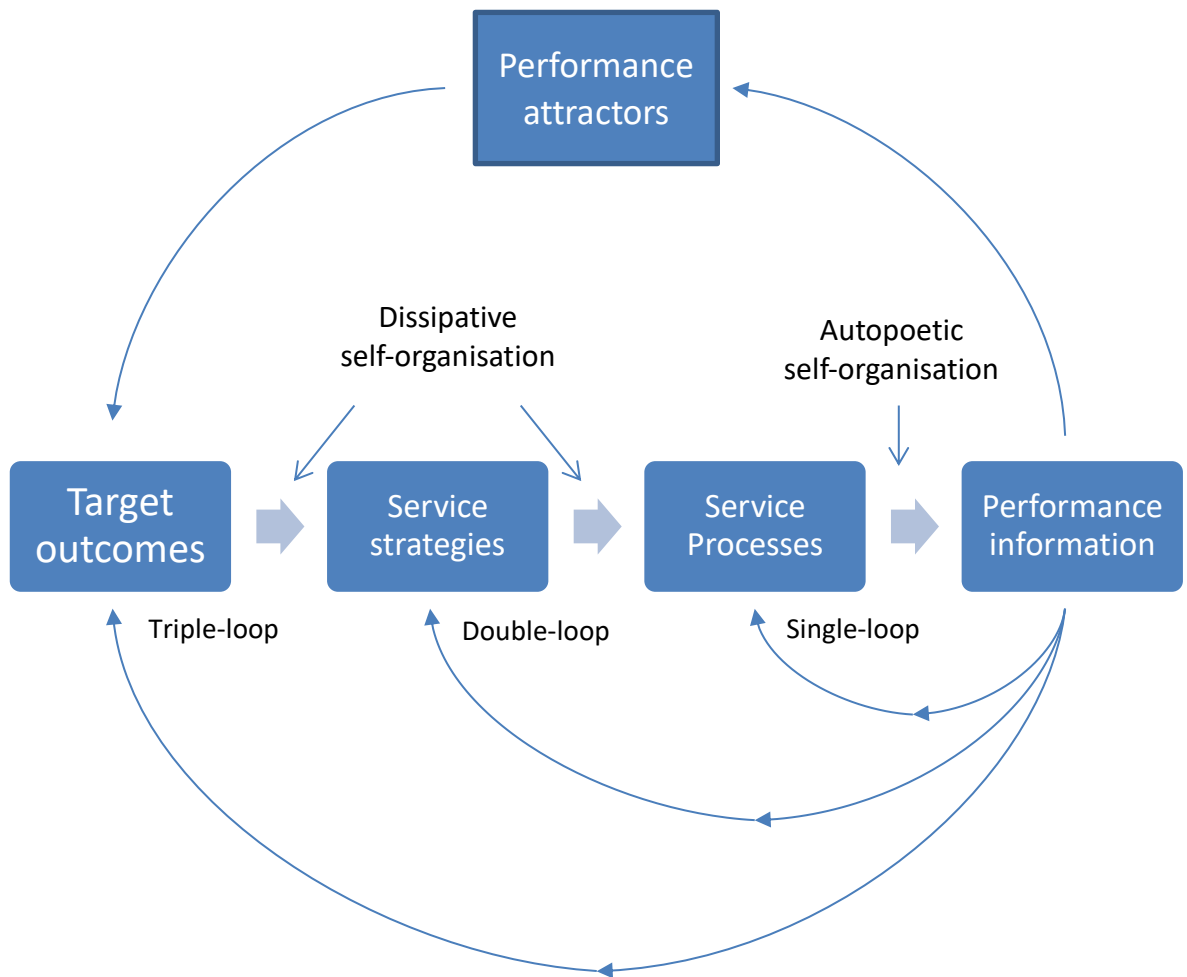
Significance of framework

This is the first theoretical model of service improvement and purposive system change which responds directly to the view of outcomes as emergent products of complex systems (Jayasinghe 2011; Lowe et al. 2016). Rather than assessing outcome-effectiveness in terms of output, its focus is on process: how the system generates and

makes sense of learning (learning capacity), how agents respond to learning by self-organising (self-organising capacity) and how the system itself can nudge self-organising dynamics of individual agents in beneficial ways relative to its best impression of outcomes (coordinative capacity). The model clarifies the task which faces outcomes-focussed service systems and social interventions within them: they must break out of an inward-focussed approach to service transformation, characterised by single-loop learning, point attractors and autopoetic self-organisation, and activate an outcome-focussed service transformation process characterised by dissipative self-organisation, strange attractors and higher order (double and triple loop) learning.

The theoretical framework's most significant contribution is to enable meaningful progress where outcomes are highly *complex*, characterised by significant causal uncertainty and ambiguity, and requiring a coordinated response among fragmented and heterarchical governance landscapes: precisely the features which have undermined the Rationalist Approach (Boyne and Law 2005; Lowe 2013; Lowe and Wilson 2015; Wimbush 2011). The goal of the Complex Systems theoretical framework is not therefore necessarily to supplant the Rationalist Approach, but to open up a class of problems which have so far proved beyond the ability of rationalist reforms to tackle.

[INSERT ANSWER TO Q4 HERE…]

The theoretical framework as an evaluative model

The Complex Systems theoretical framework provides public management scholarship with a means of moving research into outcomes-based service improvement forward on a sounder theoretical basis. To this end, Chapter 3 provides a clear set of criteria for the class of service initiatives for which the framework is suited, a coherent focus on

three interlinked adaptive capacities, and a consistent ontological and epistemological position rooted in an ontological realism and an epistemological constructivism. These developments provide surer footing for further theoretical development and more systematic empirical analysis, and can serve as a touching point for complexity-informed inquiry within public administration and public policy research more generally. The framework is not intended as an end point of theoretical development for outcomes-based approaches, and future empirical research may help to augment or refine the model and extend its explanatory power.

[OR HERE…]

Contribution to Research Question 1

The application of the theoretical framework in an evaluative capacity has focussed analysis on three areas which together constitute the system's adaptive capacity. In so doing, it has shown that where the EYC has best activated these capacities, it has done so by substantially deviating from the QIC model on which it is based.

Regarding learning capacity, the data show that the quantitative focus of the MFI was not effective of itself in generating knowledge, and that requirements of measurability promoted an undesirable focus on the more easily-measured aspects of service uptake and process, rather than an outcome-focus. Effective double-loop learning instead hinged on the adoption of a reflective qualitative approach which was driven in large part by service user co-production. Finally, the importance of a reflective planning stage was also crucial in order to facilitate triple-loop learning and the necessary path-breaking dissipative behaviour to stimulate more transformative change.

Regarding coordinative capacity, the data show the EYC failed to articulate a shared vision which was sufficiently relevant and resonant to effect dissipative behaviour

290

across its membership. Actors at the CPP-strategic and operational levels struggled to understand the relevance of Stretch Aims or KCAs for their practice, and instead self-organised in an autopoetic fashion, applying the EYC as a method to enhance the achievement of strategic priorities. Secondly, there was a lack of appreciation of agentic motivation which meant the EYC was not able to articulate performance attractors which resonated with the intrinsic motivations of the early years workforce. Thematic goals, including KCAs and associated Key Change Events which structured collaboration around shared priorities, were more effective in coordinating behaviours than age-based Stretch Aims, however were not an integral component of the EYC's structure.

Finally, regarding self-organising capacity, the data show that broadly autopoetic transformation occurred, driven by the point attractor of existing priorities, which decomposed Home Teams and led the EYC to enhance, rather than modify, existing trajectories at a CPP level. It has found that heterogeneity of service context and the 'stickiness' (von Hippel 1994) of knowledge generated led the broadly impersonal platforms for interaction (notably the Workstreams and the Extranet) to fail. Positive examples of transfer instead occurred around peer-based approaches aligned around distinct service problems.

While the Scottish Government viewed the deviation from a tried-and-tested methodology as a key factor behind its limited spread and scale (Scottish Government 2014), the application of the framework outlines a different explanation: a failure to understand factors driving self-organising behaviours and capitalise on emergent trends as the QIC model interacted with a vastly different service environment.

In line with the study's critical realist position, the conceptual power of the framework is demonstrated by its ability to distinguish between problems of implementation and underlying mechanisms of change. These are explained as attractor states, self-organising potentials, and forms of learning generation, which interlink to provide an

impression of the EYC's adaptive capacity. This allows findings to be understood not just of the EYC as a distinctive entity, but as an exemplar of the QIC method.

Limitations and suggestions for further research

The achievement of Research Objective 3 and satisfaction of Research Question 1 is marked by a number of limitations however. On a methodological level, the small sample size carries with it a danger of misrepresentation of the wider experience across the 32 CPPs. This risk is managed however through a purposive sampling strategy which ensures the CPPs are conceptually representative of the wider experience, and key characteristics (e.g. number of TOCs, shared experience of key events such as collapse of the Workstreams) are cross-checked to accord with the wider EYC experience.

Perhaps more significantly, the empirical analysis in Chapters 5 and 6 raises many issues which are not directly addressed by the theoretical framework. Issues of power were key in constraining the EYC's observed development, yet the framework does not address in significant depth the intricacies of power relationships in organisational settings. Similarly, institutional factors, including organisational culture and complex accountability structures were observed to constrain the opportunity space both for CPP strategic direction and TOC emergence and development. Cultural and institutional factors are noted to be crucial to enabling change in complex systems (Argyris and Schön 1974; Sterman 1994), however again this area has no explicit role within the framework. Finally, the role of leadership in complex systems (Lichtenstein and Plowman 2009; Schneider and Somers 2006; Uhl-Bien et al. 2007), though debated through the discussion of performance attractors, also has no formal role.

In lacking an explicit place for crucial issues of power, institutional factors, or leadership, the framework as an evaluative model might be accused of presenting an overly simplistic view of change within complex systems. The theoretical framework should therefore not be assumed to be a comprehensive explanation of system

transformation, which requires the confluence of an abundance of factors, some of which Koch et al. (2016, 291) note: 'connectors and integrators that span the boundaries, sharing of goals among participants, aligned funding and incentives, and a supporting infrastructure, all leading to a virtuous cycle of collaboration'.

The application of the Complex Systems theoretical framework is nevertheless conceptually congruent, despite these simplifications, in drawing inference about the key mechanisms and contexts which have vitiated or enabled adaptive behaviour, and reveals progress that would go unappreciated if a linear programme theory evaluation were applied, incidentally demonstrated by the Scottish Government's (2014) Stock Take evaluation. While recognising the validity of criticisms of the model's simplicity, the thesis nevertheless argues that a simplified framework has value as a conceptual heuristic, the clarity of which would be diminished by further accumulation of cultural or institutional factors, or issues of leadership and accountability. Any theoretical framework will fall short of a comprehensive explanation of change in complex system, wherein explanatory and predictive power is redefined as the potential to simplify complexity while still remaining congruent with real causative events (Jessop 1997).

Contribution to Research Question 2

Co-production research has been one of the main growth areas of public administration research over the last decade (Osborne 2016). RQ 2 responds to a growing call from research and policymakers for more clarity over how the creative potential of co-production can be harnessed for service improvement (Bovaird et al. 2017; Bovaird and Loeffler 2016; Voorberg et al. 2014). Chapter 7 provides rare micro-level empirical evidence which ties forms of co-production to discrete outcomes.

Drawing on data from TOC case studies and the experience of the Locality Model, Chapter 7 adapts the Complex Systems theoretical model to provide a novel framework for understanding how learning and improvement can be generated

through co-production, presented again for convenience in Figure 11. This model clarifies how service user feedback can improve the efficiency (inputs-outputs), effectiveness (outputs-outcomes) and relevance (redefining outcomes) of public services through differential pathways. Taken together, the model provides a framework whereby the transformational potential of co-production, outlined often in co-production scholarship but rarely seen in practice might be operationalised (Cahn and Grey 2012; Needham and Carr 2008; Osborne and Strokosch 2013).

Figure 11. The co-production of learning
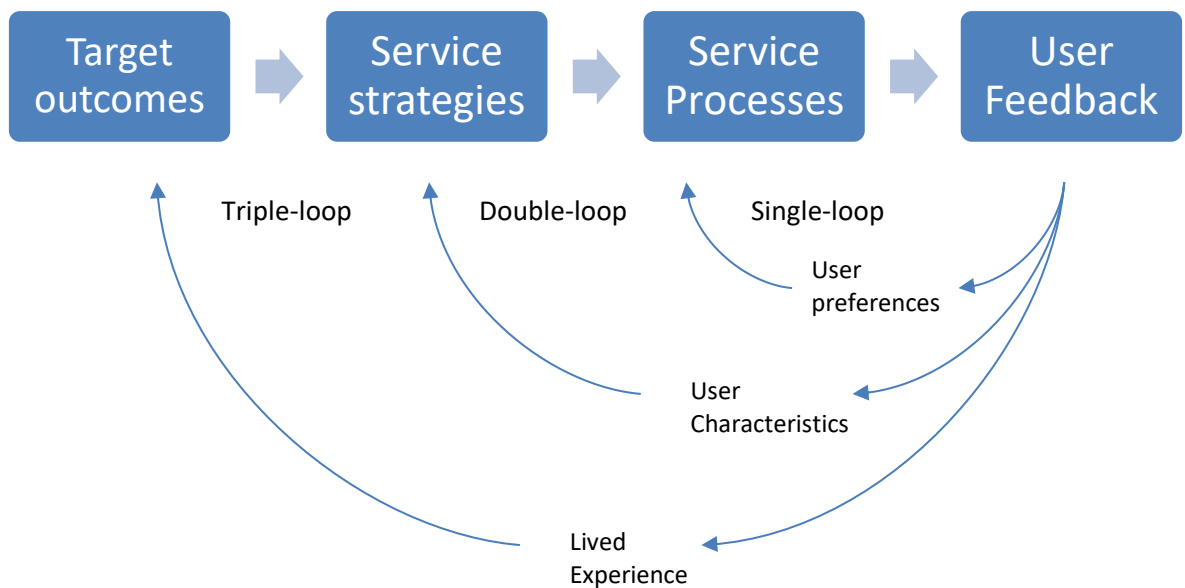
Contributions to public management research

Echoing literature drawing from services marketing (Chen et al. 2011; Osbourne and Radnor 2015), the model suggests service users and communities of identity (i.e. those experiencing outcomes) have a central role within service improvement and innovation, however it considers that not only service experiences, preferences and characteristics (Chen et al. 2011), but the wider lived experience of these individuals (van Manen 1990) can serve as inputs into specifically outcomes-focussed service transformation. While the creative potential of co-production is often linked to service innovation (e.g. Voorberg et al. 2014), the contribution of service users to bettering public services through their contribution to the micro-processes of learning has not so far entered this debate. The model clarifies the importance of learning to individual

and organisational transformation, showing how change at the organisational and system level can be informed through a bottom-up feedback process.

For co-production scholarship, the model provides a more solid basis for fine-tuned and detailed inquiry into the creative potential of co-production, including future scholarship into public service innovation, social innovation, user-led improvement, co-creation and service design. For policymakers, service commissioners and public managers, the model provides further clarity over the expected risks and rewards of incorporating different models of co-production, allowing more informed choices to be made regarding strategic management and programme design. A practical guide to support decision making about how and when to incorporate co-production into organisational design is presented in Table 14 in Chapter 7.

Bovaird et al. (2015) suggest that the contribution of service users is greatly underutilised within contemporary services, particularly given its potential within a time of fiscal constraint. Findings accord with this assertion, however suggest a similarly significant problem is that the already considerable contributions of co-production are often unrecognised. While user feedback led to the most important innovations achieved in TOC case studies, this took place through tacit forms of feedback which took place in more informal and organic methods of user involvement – observation, conversation and discussion, focussed at the practitioner-user level where trusting relationships exist – rather than traditional instruments of involvement (e.g. consultations, focus groups, forms of representation). Maximising the benefits of co-production for public value and better services therefore likely requires a conceptual shift to make co-production explicit and encourage more systematic harnessing of user input.

Finally, and most significantly for the research aim, the framework clarifies the role of co-production in complex systems change, showing how co-production can be harnessed to improve outcomes as well as service efficiency. The analysis suggests

that co-production can serve as a key source of learning generation in complex systems, which the Complex Systems theoretical framework suggests can facilitate outcomes-focussed system transformation.

Limitations and suggestions for further research

While the thesis contends that a focus on learning is significant, a learning perspective can only be a partial explanation of co-production's potential impact. Many throughputs – contravening strategic priorities, resource constraints, cultural barriers, and poor connectedness of practitioners – have conspired to assure that improvements realised in the data lagged significantly behind learning. This thesis stops short of discussing the relationship between public service innovation (broadly, the enactment of novel and disruptive ways of working) and public service learning, a much broader area of research necessary but not sufficient for more fully understanding the creative potential of co-production.

The learning framework by itself also does not specifically address how co-production of learning can be institutionalised, which is also a worthy avenue for further explanation. In practice, there are many tools which might fit this purpose: QI may be hold potential if optimised to take account of co-production, while service design methodologies also hold potential to advance co-production of learning. Work remains to be done integrating such practical methods with a theoretically consistent learning framework such as that presented here.

Contribution to Research Question 3

The proliferation of QICs as population-outcome focussed interventions has given rise to a growing body of literature, most of which has advocated the model (Inkelas and Bowie 2014; Bryk et al. 2011; Ghandour et al. 2017; Green et al. 2012; Inkelas and McPherson 2015; McPherson et al. 2015). However, this body of literature lacks significant in-depth empirical work, and the significant conceptual differences concerning population-level and clinical outcomes normally targeted by QICs leaves a

substantial lacuna within this scholarship. This chapter contributes the first substantial empirical study of a QIC operating in a multi-agency context to achieve a defined set of population-level outcomes, and is thus in a privileged position to inform this emerging area of scholarship.

Population-level outcomes present a different challenge to QICs, as while indicators of institutional quality are produced directly through the interaction of patients with the healthcare system, the emergence of thematically similar population outcomes are determined predominantly by interactions outwith the institutional setting. Achieving population-level outcomes therefore pits QI initiatives against a far more substantial challenge, which the empirical analysis suggests surfaces three particular tensions:

- The social world is innately less measurable than clinical healthcare settings and impact, particularly regarding high-level outcomes, is difficult to appreciate through quantitative data alone.
- Social service systems are more loosely connected and heterarchical than clinical settings, with significant autonomy and the presence of multiple intersecting and contradicting goals.
- Service fragmentation, contextual uniqueness and heterogeneity of service delivery contexts makes knowledge stickier (von Hippel 1994) and less easily transferred compared to clinical settings.

Learning in the multi-agency environments which characterise population-outcome systems was achieved not as a technical process of evidencing and transference, but as a social process of participation, deliberation and peer collaboration (Wenger 1998). The findings also suggest that a broader and more methodologically inclusive approach to QI is needed than is commonly espoused by its advocates within the healthcare sector. The technical and quantitative focus of the MFI was only appropriate in a relatively small subset of service contexts, and was poorly equipped to determine service impact.

Chapter 8 suggests a number of ways in which the aforementioned challenges can be overcome. To improve the system's learning capacity, QICs could permit a 'pipeline' approach to TOC development, allowing more rapid and non-linear development during early stages, and relaxing methodological rigour assumptions until the projects have achieved maturity, similar to approach taken by Project Oracle (Ilic 2011), and within Nesta's standards of evidence (Puttick and Ludlow 2013). Secondly, QICs could carve out an explicit role for co-production at the 'plan' and 'study' stage during testing to enhance the innovative potential of TOCs. To improve coordinative capacity, QICs could insist on an orientation stage prior to testing in explicit reference to performance attractors, perhaps drawing on design thinking tools. QICs may also benefit by developing an evolving understanding of intrinsic motivations of actors to identify opportunities for nudging behaviours in collectively beneficial directions. Finally, QICs could ensure to build a more inclusive shared vision which is understood and valued among key actors, which might be achieved through a deliberative system mapping exercise.

Finally, to improve self-organising capacity, QICs could more explicitly focus on peer-led collaboration around shared goals, perhaps through explicit Communities of Practice organised around specific issues (Wenger 1998). The diverse and fragmented service context mean that knowledge must be mobilised across diverse contexts, and thus require greater integrative adaptivity (Bryk et al. 2016), which might be facilitated through configuring the PM role as active boundary spanners with responsibility for adapting learning across diverse contexts.

Limitations and suggestions for further research

This is the first substantial empirical study of QICs operating in a population outcome setting. Kennedy (1979) argues that for reasonable generalisation from a single case, a wide range of attributes must be present across both the case and wider population, many of which are shared and few of which are unique to the individual case. The extent to which the attributes uncovered in this analysis are shared is unclear, since

we lack comparable detailed and specified case studies of other outcomes-focussed QICs.

While the critical realist position of this research allows for the theoretical potential of analytical generalisation through its consideration of an objective reality, its epistemological constructivism, reflecting the inherent boundedness of situated agents operating within complex systems, diminishes the likelihood of universal theories being generated from single sites. Research in this area has been inhibited by the paucity of initiatives operating at a large scale. However, with the emergence of the population-level QICs in Scotland and the US, there exists a growing body of initiatives from which to undertake more theoretically-informed comparative research.

Moving forward with a Complex Systems Approach to public service improvement: a tentative research agenda

The development of the theoretical framework is intended not just to apply to QIC initiatives, but to support outcomes-focussed interventions and programmes of service reform more generally. Accordingly, Research Objective 4 commits to a consideration of how inquiry into the design of outcomes-focussed reforms and interventions might proceed. This final section reflects on thesis findings to suggest some routes forward for future research. Two avenues of inquiry are suggested which might be taken further in public management and practice: applying the framework to other interventions such as Communities of Practice (Wenger 1998) and Parallel Organisations (Zand 1974), and applying the theory in a developmental capacity through Developmental Evaluation (Patton 1994). Each avenue is briefly explored in the following section.

Applying the framework to other interventions

The application of the framework to the case of the EYC brings up two key issues which inhibited progress: the difficulty of 'attracting' improvement focus to system outcomes, rather than organisational goals, and the difficulty of communicating and 'scaling' learning within the system. It is not possible to know from this application how general these issues are, or if there are many other common issues which the EYC happens to have gotten right. The three criteria which interventions are expected to meet for applying the model (themselves also open to reappraisal) provide future scholarship with clear guidance for adding to this body of research on a more systematic and comparable basis. The theoretical model is well placed to facilitate comparative research on the fitness of outcomes-focussed systems, where a best practice is as Bryk et al. (2011, p.34) argue, 'one that grows and sustains participation, focuses ongoing efforts on targeted priorities, and ultimately contributes to improvement reliably at scale'.

Communities of Practice

Two particular interventions have strong conceptual potential for more systematically applying the theoretical framework. Chapter 8 has noted the similarities between the QIC model and Wenger's (1998) Communities of Practice. Lowe et al. (2016) have outlined Communities of Practice as a potential route to a complexity-friendly performance management through facilitating horizontal accountability and a 'positive error culture' (Gigerenzer 2015), leading to better collaborative innovation in conditions of uncertainty.

Communities of Practice are based on the understanding of learning as a social construct which is shaped through continuous participation without distinct beginning or end points (Wenger 1998). This stands in contrast with the BSC model's technical understanding of learning as a mechanical process which, when evidence is generated, can be packaged and transported across boundaries. The social understanding of learning better fits the experience of sharing and spreading learning within the EYC. By focussing on significant service problems, Communities of Practice may promote

better focussed innovation around key leverage points (such as the Key Change Areas in the case of the EYC), and provide a potential route to more productive outcomes-focussed system transformation. This line of inquiry has potential to better mobilise knowledge which is 'sticky' at boundaries (Dopson and Fitzgerald 2005) and less straightforwardly transferred across diverse contexts.

However, Communities of Practice are not common at a large scale, nor on a multi-agency basis. Moving in this direction problematises the concept since the recognition of competency essential for the development of horizontal accountability is more difficult where actors have less understanding of one another's practice. One route forward is to encourage sharing across boundaries through a dedicated boundary spanning role to respond actively and opportunistically to spread learning between problem-focussed Communities of Practice. This could activate what Bryk et al. (2011) term 'integrative' adaptivity, mobilising learning across boundaries where self-organisation would not take place organically.

This requires any system-level improvement community to be 'constituted by interrelated communities of practice' (Wenger 2010, p.1). Linking together problem-focussed Communities of Practice is a potential means of resolving the relative limitations of an improvement 'community' (Englebart 1992) in which the patterns of interactions conform to stability and predictability, and an improvement 'network' (Bryk et al. 2011) which can be fragmented and poorly focussed. In a complex system, this pattern of lateral and vertical integration conforms to a situation of 'bounded instability' (Merry 1999; Stacey 1995), in which 'the organisation can find the mix of confirmation and novelty that allows it to be a learning system that is able continually to self-organize and thus renew itself' (Merry 1999, p.275).

Parallel Organisations

The EYC lacked an explicit approach to creating conducive institutional environments, expecting that innovation towards system goals would occur alongside existing

organisational functionality. However, the data suggest that this approach failed to instigate significant redirection in purpose of improvement focus. Organisational theory has engaged similarly with how to accommodate complex, cross-cutting and ill-defined problems alongside organisational functions. Sanderson (2000) for instance suggests creatiing 'local flora', or sanctuaries from coercive power structures, while Senge's (1990) emphasis on creative problem solving teams similarly aims to encourage local exploration independent of institutional hierarchy. The increasing prominence of design thinking within management has brought focus on incorporating a diversity of perspectives for problem solving (Gruber et al. 2015).

Hawk and Zand (2013) argue that poorly defined and cross-cutting problems require a bimodal operation to facilitate an 'organisational ambidexterity' (Raisch and Birkinshaw 2008), dually focussed on internal efficiency and external effectiveness. A promising approach in this vein is the 'collateral' or 'parallel' organisation model developed by Zand (1974), and later expounded by Kilmann (1982) and Hawk and Zand (2013). Zand's (1974) approach involves the creation of a formal problem-focussed parallel organisation structure alongside and separate to organisational processes conducted within the operational organisation. Zand (1974) recommends that team members within the parallel organisation are committed to spend 2-10 hours per week working on common problems. The overlapping membership of the parallel and operational organisations allows problem sensing and solution design to be linked with implementation within vertically accountable organisational teams.

There are examples of parallel organisations operating at a large scale. Hawk and Zand (2013) for instance report a successful case involving 360 managers within a large organisation. However there are no examples of parallel organisations operating in a public sector context, and no discussion of the fit of the model within a system (extra-organisational and multi-agency) level. Future research could look for ways to formally adapt the parallel organisation model to a system level. Kilmann (1982) for instance

outlines a ten step process for designing collateral organisations, which could be updated to function at a system level.

Developmental evaluation

The application of the theoretical model in an evaluative capacity has natural limitations, given that the framework is not consciously applied and its prescriptions are not followed. Thus the application of the theoretical framework in a diagnostic fashion cannot verify its prescriptive power. This means that the potential for the framework as a conscious tool of policy design and improvement is untested.

Complexity-consistent approaches to evaluation, such as Outcome Mapping, Outcomes Harvesting and Developmental Evaluation may serve to integrate the framework with practice. These are all approaches which are specifically designed for operation in settings where strategic planning is limited and action must follow learning more rapidly than traditional programme evaluation cycles.

Of the three approaches, Developmental Evaluation has the most significant tradition in complex systems research. Developmental Evaluation aims to 'support innovation within a context of uncertainty' (Gamble 2008 p.15) through an embedded, continuous process of sense-making and adaptation. It is suited for adopting in ambitious long-term approaches within development and social programmes focussed on soft issues which are difficult to measure or even to understand when they have been achieved. Developmental Evaluation reframes the role of evaluator as facilitator and bricoleur (Dozois et al. 2010), engaged in sense-making about shared problems (Patton 1994).

Conclusions: paths forward for public management

The application of the Complex Systems theoretical model in an evaluative capacity has natural limitations, since it is not consciously applied and cannot influence decisions made on a dynamic basis. Thus a final contribution made here, in fulfilment

of Research Objective 4, has been to consider how the framework can be better integrated with practice. Three potential routes are considered here: Communities of Practice (Wenger 1998), Parallel Organisations (Zand 1974), and Developmental Evaluation (Patton 1994). Table 16 below sets these three approaches against the theoretical framework's criteria for adoption set out in Chapter 3, showing that each is a good conceptual fit. This provides future scholarship in extending a Complex Systems approach with potential inroads to testing the theoretical framework's value as a practical model. The resonance of the theoretical framework with these isolated and diverse approaches also illustrates the potential of the framework as a broader paradigmatic alternative to the Rationalist Approach. Further research could continue to build an ensemble of approaches - to commissioning, institutional design, evaluation and monitoring, and planning and performance management - to rival their Rationalist counterparts.

Table 16. Opportunities for extending research into practice

Condition for application of theoretical framework	Communities of Practice	Parallel Organisations	Developmental Evaluation
Systems must have articulated a clear outcomes-focus, but agents must lack formal accountability for their achievement	Communities of Practice are arranged around shared goals considered to be primary drivers of outcomes – accountability is horizontal, shared across peers	Parallel Organisations are similarly focused around shared goals, however learning is linked to organisational accountabilities in the operating organisation	Evaluators can focus teams on a shared vision, generate and maintain a sense of direction, and help to understand system dynamics (Dozois et al. 2010). Team members are not linked by formal accountability, rather evaluation is focussed on generating learning
Agents must possess significant autonomy over how they choose to organise	Agents must voluntarily join Communities of Practice, however self-organisation is limited to similar settings – boundary spanning across problem-focussed Communities of Practice needed at the system level	Effectiveness depends on voluntary commitment to parallel infrastructure and resultant self-organising responses	Evaluators can actively facilitate the process of self-organisation through developing relationships, building connections and responding to emergent learning
Agents must be capable of generating and transmitting feedback from actions to other agents, creating systemic knowledge	Learning is achieved through participation and indirect peer-led influence. Learning between diverse contexts can be actively facilitated through boundary spanning	Lacks a specific method of generating learning. The organisational and parallel structures must be engaged in constant feedback between learning and practice	Evaluators ensure that learning is understood and reflected on in team meetings and throughout organisational processes

Concluding summary: what potential for an outcomes-based approach to public managment?

This thesis' Research Aim has been to improve the ability of public service systems to achieve valued outcomes. It has sought to achieve this through improving conceptual understanding of the thesis' subject of outcomes (Research Objective 1), developing an appropriate alternative theoretical approach and grounding this empirically (Research Objectives 2 and 3 respectively) and finally through considering the implications of the findings for future research and practice (Research Objective 4).

This thesis has attempted to demonstrate that outcomes are an important concept in public administration, public policy and management with implications beyond the technical issues they are traditionally understood to present within such literatures (Heinrich 2002; Boyne and Law 2005; Schedler and Proeller 2010). It argues that outcomes should be considered emergent products of complex systems, with fundamental challenges to the way that public services are organised and delivered. Conversant with emerging literature in public administration (Lowe and Wilson 2016; Lowe et al. 2016; Lowe 2017), it argues complexity theory holds particular promise not just in conceptualising outcomes, but in developing an alternative Complex Systems Approach to outcomes-based management which transcends the limitations of the NPM-derived Rationalist Approach. The theoretical framework developed in Chapter 3 demonstrates conceptual validity through elucidating many of the mechanisms driving both success and failure within the Early Years Collaborative, many of which lie contrary to the Breakthrough Series Collaborative model on which the initiative is based. The framework's evaluative power also allows insights to be drawn against two specific Research Questions – relating to the potential impacts of user co-production on service improvement, and the potential application of QICs in population-focussed service systems, showing its potential to add novel theoretical insight to contemporary public management problems.

However, the thesis cannot by itself vindicate a Complex Systems approach to outcomes management. The Early Years Collaborative, three years following its initiation, had achieved far more modest results that the Scottish Patient Safety Programme on which its structure and methodology was based. Before concluding therefore, it is worth entertaining a thought experiment: what might have been expected if Community Planning Partnerships were made directly accountable for delivery against the Stretch Aims, as within the Rationalist Approach?

Take Stretch Aim 2 (Appendix H) for instance, which aimed to improve to 85% the proportion of children achieving developmental goals across physical, social, behavioural and developmental domains at the population level within localities. In this case, the public management literature reviewed would predict a redirection of resources to children failing just a small number of barriers, particularly those areas (e.g. areas of speech and language, rather than parental attachment) which locality managers felt could be more predictably influenced. We might also reasonably expect to encounter 'creaming' and 'parking' behaviour (van Thiel and Leeuw 2002), with far greater tolerance of the observed tendency for many at-risk families to fail to put themselves forward for review – we would almost certainly not have seen the significant impetus across Scotland to improve uptake for precisely this group. While this analysis may appear cynical, as Chapter 2 documents in greater depth, similar perverse behaviours are demonstrated at length, systematically, within the public management literature.

In lieu of this alternative, the theoretical framework provides a promising starting point to begin to articulate how workable alternatives to the improvement of outcomes may be developed in way which transcends the demands of the NPM-linked Rationalist Approach. Three such approaches are suggested in this final chapter, alongside the modified Quality Improvement Collaborative structure suggested in Chapter 8. These include incorporating the framework within institutional design and strategic management through Developmental Evaluation (Patton 1994), and

extending the framework in practice through the application of Zand's (1974) Parallel Organisations or Wenger's (1998) Communities of Practice.

The thesis introduction has posed the question, *how can public service outcomes be sustained or improved in the face of worsening demand and stagnating levels of input?* It is ultimately hoped that the contributions of the thesis detailed in this chapter can inform the development of functional alternatives to New Public Management in the pursuit of viable solutions to this increasingly problem.

References

Ackroyd, S. (2004) Some implications of critical realism. *Critical Realist Applications in Organisation and Management Studies*, 11, pp. 127.

Ackroyd, S. and Fleetwood, S. (2000) *Realist perspectives on management and organisations*. Hove: Psychology Press.

Agee, J. (2009) Developing qualitative research questions: a reflective process. *International Journal of Qualitative Studies in Education*, 22(4), pp. 431-447.

Ames, N. (1999). Social work recording: A new look at an old issue. *Journal of Social Work Education*, *35*(2), pp. 227-237.

Anaf, S. Drummond, C. and Sheppard, L.A. (2007) Combining case study research and systems theory as a heuristic model. *Qualitative Health Research*, 17(10), pp. 1309-1315.

Anderson, P., Meyer, A., Eisenhardt, K., Carley, K. and Pettigrew, A. (1999) Introduction to the special issue: Applications of complexity theory to organization science. *Organization Science*, 10(3), pp. 233-236.

Anderson, R.A., Crabtree, B.F., Steele, D.J. and McDaniel Jr, R.R. (2005) Case study research: the view from complexity science. *Qualitative Health Research*, 15(5), pp. 669-685.

Argyris, C. (1977) Double loop learning in organizations. *Harvard Business Review* pp. 115–125.

Argyris, C. and Schön, D.A. (1974) *Theory in practice: Increasing professional effectiveness*. San Francisco: Jossey-Bass.

Argyris, C., Schön, D. A. (1978) *Organizational learning*. Reading: Addison-Wesley.

Arnstein, S.R. (1969) A ladder of citizen participation. *Journal of the American Institute of planners*, 35(4), pp. 216-224.

Ashby, W.R. (1956) An introduction to cybernetics. *An introduction to cybernetics. London:* Chapman & Hall.

Audit Commission (2000) *On target: the practice of performance indicators.* Audit Commission, London.

Audit Scotland (2014) *Self-directed support.* Edinburgh: Audit Scotland.

Axelrod, R. and Cohen, M.D. (2000) *Harnessing complexity: Organizational implications of a scientific frontier*. Basic Books.

Ball, S., Mudd, J., Oxley, M., Pinnock, M., Qureshi, H. and Nicholas, E. (2004) Make outcomes your big idea: using outcomes to refocus social care practice and information. *Journal of Integrated Care*, 12(5), pp. 13-19.

Barnes, M., Matka, E. and Sullivan, H. (2003) Evidence, understanding and complexity: evaluation in non-linear systems. *Evaluation*, 9(3), pp. 265-284.

Batalden, P.B. and Davidoff, F. (2007) What is "quality improvement" and how can it transform healthcare? *Quality and Safety in Health Care*, 16, pp. 2–3.

Bate, P. and Robert, G. (2007) *Bringing user experience to healthcare improvement: The concepts, methods and practices of experience-based design*. Abingdon: Radcliffe Publishing.

Baxter, P. and Jack, S. (2008) Qualitative case study methodology: Study design and implementation for novice researchers. *The qualitative report*, 13(4), pp. 544-559.

Beitsch, L.M., Thielen, L., Mays, G., Brewer, R.A., Kimbrell, J., Chang, C., Gillen, S., Corso, L. and Landrum, L.B. (2006) The multistate learning collaborative, states as laboratories: informing the national public health accreditation dialogue. *Journal of Public Health Management and Practice*, 12(3), pp. 217-231.

Beer, S. (1984) The viable system model: Its provenance, development, methodology and pathology. *Journal of the operational research society*, 35(1), pp. 7-25.

Berwick, D.M., Nolan, T.W. and Whittington, J. (2008) The triple aim: care, health, and cost. *Health affairs*, 27(3), pp. 759-769.

Beresford, P. and Boxall, K. (2013) Where do service users' knowledges sit in relation to professional and academic understandings of knowledge? In: *Mental Health Service Users in Research: Critical Sociological Perspectives*, pp. 69-86.

Berkes, F., Colding, J. and Folke, C. (2003) *Navigating social–ecological systems: building resilience for complexity and change.* Cambridge: Cambridge University Press.

Bevan, G. and Hamblin, R. (2009) Hitting and missing targets by ambulance services for emergency calls: effects of different systems of performance measurement within the UK. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(1), pp. 161-190.

Bevan, G. and Hood, C. (2006) What's measured is what matters: targets and gaming in the English public health care system. *Public administration*, 84(3), pp. 517-538.

Bevan, G. and Wilson, D. (2013) Does 'naming and shaming' work for schools and hospitals? Lessons from natural experiments following devolution in England and Wales. *Public Money & Management*, 33(4), pp. 245-252.

Bevir, M. and Rhodes, R.A. (2012) Interpretivism and the analysis of traditions and practices. *Critical policy studies*, 6(2), pp. 201-208.

Bhaskar, R. (2010) *Reclaiming reality: A critical introduction to contemporary philosophy*. London: Taylor & Francis.

Bhaskar, R. (2013) *A realist theory of science*. London: Routledge.

Bianchi, C. (2016) *Dynamic Performance Management*. New York: Springer International Publishing.

Birchall, J. and Simmons, R. (2004) What Motivates Members to Participate in Co-operative and Mutual Businesses? *Annals of Public and Cooperative Economics*, 75(3), pp. 465-495.

Blackman, T. (2013) Rethinking policy-related research: charting a path using qualitative comparative analysis and complexity theory. *Contemporary Social Science*, 8(3), pp. 333-345.

Blair, T. (1998) *Leading the Way: A New Vision for Local Government*. London, IPPR.

Bogason Peter. (2005) *Postmodern public administration. In Oxford handbook of public management, ed. Ferlie Ewan Lynn Laurence E.* Pollitt Christopher 234–56 Oxford, UK: Oxford Univ. Press.

Booher, D.E. and Innes, J.E. (2010) Governance for resilience: CALFED as a complex adaptive network for resource management. *Ecology and Society*, 15(3), p.35.

Bovaird, T. (2007) Beyond engagement and participation: User and community coproduction of public services. *Public administration review*, 67(5), pp. 846-860.

Bovaird, T. (2008) Emergent strategic management and planning mechanisms in complex adaptive systems: the case of the UK Best Value initiative. *Public Management Review*, 10(3), 319-340.

Bovaird, T. (2010) *Model of public outcomes and cost implications: Thinkpiece for Birmingham total place initiative*. Birmingham: INLOGOV.

Bovaird, T. and Loeffler, E. (2013) We're all in this together: harnessing user and community co-production of public outcomes. *Birmingham: Institute of Local Government Studies.*

Bovaird, T. (2014) Attributing Outcomes to Social Policy Interventions–'Gold Standard'or 'Fool's Gold'in Public Policy and Management? *Social Policy & Administration*, 48(1), 1-23.

Bovaird, T. and Loeffler, E. (2016) Bringing the Resources of Citizens into Public Governance. In Torfing, J. and Triantafillou, P. eds. *Enhancing Public Innovation by Transforming Public Governance*. Cambridge: Cambridge University Press.

Bovaird, T., Flemig, S., Loeffler, E. and Osborne, S.P. (2017) Debate: Co-production of public services and outcomes. *Public Money & Management*, pp. 1-2.

Bowen, G.A. (2009) Document analysis as a qualitative research method. *Qualitative research journal*, 9(2), pp. 27-40.

Boland, T. and Fowler, A. (2000) A systems perspective of performance management in public sector organisations. *International Journal of Public Sector Management*, 13(5), pp. 417-446.

Boyne, G.A. and Law, J. (2005) Setting public service outcome targets: lessons from local public service agreements. *Public Money & Management*, 25(4), pp. 253-260.

Broadie, A. (2010) *The Scottish Enlightenment* (Vol. 80). Edinburgh: Canongate Books.

Bromley, D. B. (1990) Academic contributions to psychological counselling: a philosophy of science for the study of individual cases. *Counseling Psychology Quarterly*, 3, 299-307.

Bryk, A.S., Gomez, L.M. and Grunow, A. (2011) Getting ideas into action: Building networked improvement communities in education. In *Frontiers in sociology of education* (pp. 127-162). Dordrecht: Springer Netherlands.

Bryman, A. (2007) The research question in social research: what is its role?. *International Journal of Social Research Methodology*, 10(1), pp. 5-20.

Bryson, J.M., Crosby, B.C. and Bloomberg, L. (2014) Public value governance: Moving beyond traditional public administration and the new public management. *Public Administration Review*, 74(4), pp. 445-456.

Buchanan, J.M. and Tullock, G. (1962) *The calculus of consent* (Vol. 3). Ann Arbor: University of Michigan Press

Burns, H. (2015) Health inequalities–why so little progress? *Public health*, 129(7), pp. 849-853.

Butler, M.J. and Allen, P.M. (2008) Understanding policy implementation processes as self-organizing systems. *Public management review*, 10(3), pp. 421-440.

Byrne, D.S. (1998) *Complexity theory and the social sciences: an introduction*. Hove: Psychology Press.

Byrne, D.S. (2011) Exploring organizational effectiveness: The value of complex realism as a frame of reference and systematic comparison as a method. *The Sage handbook of complexity and management*, pp. 131-141.

Byrne, D. and Uprichard, E. (2012) Useful complex causality. *The Oxford handbook of philosophy of social science*, pp. 109-29.

Byrne, D. and Callaghan, G. (2013) *Complexity theory and the social sciences: The state of the art*. Routledge.

Bywaters, P. Brady, G. Sparks, T. and Bos, E. 2014. Child welfare inequalities: New evidence, further questions. *Child & Family Social Work*, 21(3), pp. 369-380

Cabinet Office (1999) *Modernising government Stationery Office*: London

Cahn, E., & Gray, C. (2012) *Co-production from a normative perspective* (pp. 129-144). V. Pestoff, T. Brandsen, & B. Verschuere (Eds.). New York, NY: Routledge.

Capra, F. (2002) *The hidden connections: integrating the biological, cognitive, and social dimensions of life into a science of sustainability.* Doubleday, New York, New York, USA.

Cairney, P. (2015) Scotland's Future Political System. *The Political Quarterly*, 86(2), pp. 217-225.

Cairney, P. (2016) Evidence-based best practice is more political than it looks: a case study of the'Scottish Approach'(Forthcoming/Available Online). *Evidence and Policy: A Journal of Research, Debate and Practice*.

Cameron, D. (2010) *Speech to Lord Mayor's banquet*. London.

Capra, F. (2002) Complexity and life. *Emergence*, 4(1-2), pp. 15-33.

Carr, S. (2011) Enabling risk and ensuring safety: self-directed support and personal budgets. *The Journal of Adult Protection*, 13(3), pp. 122-136.

Carr, W. and Kemmis, S. (1983) *Becoming critical: Knowing through action research*. Burwood: Deakin University.

Cartwright, Nancy, Bradburn, Norman M. & Fuller, Jonathan (2016) *A Theory of Measurement*. Working Paper: 2016-07. Durham: Centre for Humanities Engaging Science and Society (CHESS).

Checkland, P. (1981) *Systems Thinking, Systems Practice*. Chichester: Wiley.

Chen, J.S., Tsou, H.T. and Ching, R.K. (2011) Co-production and its effects on service innovation. *Industrial Marketing Management*, 40(8), pp. 1331-1346.

Chief Medical Officer for Scotland (2012) *Annual Report of the Chief Medical Officer 2012: Population Health and Improvement Science*. Edinburgh: Scottish Government.

Children and Families Analysis (2014) *Key Changes – Evidence Base with Evidence Sources*. Edinburgh: Scottish Government.

Chiva-Gómez, R. (2003) The facilitating factors for organizational learning: bringing ideas from complex adaptive systems. *Knowledge and Process Management*, 10(2), pp. 99-114.

Christensen, T. and Lægreid, P. (2007) The whole-of-government approach to public sector reform. *Public administration review*, 67(6), pp. 1059-1066.

Christie Commission. (2011) *Report on the Future Delivery of Public Services*. Edinburgh: Scottish Government.

Cilliers, P. (2001) Boundaries, hierarchies and networks in complex systems. *International Journal of Innovation Management, 5, pp.* 135–147.

CIPFA Scotland (2013) *Public finances: at the edge of chaos and ready for outcomes?* Edinburgh: CIPFA.

Clancy, C.M. and Eisenberg, J.M. (1998) Outcomes research: measuring the end results of health care. *Science*, *282*(5387), pp. 245-246.

Cochran-Smith, M., Ell, F., Grudnoff, L., Ludlow, L., Haigh, M. and Hill, M. (2014) When Complexity Theory Meets Critical Realism: A Platform for Research. *Teacher Education Quarterly*, 41(1), pp. 105-122.

Comfort, L. K. (1994) Self-organization in complex systems. *Journal of Public Administration Research and Theory*, 4(3), 393–410.

Connick, S., and J. E. Innes (2003) Outcomes of collaborative water policy making: applying

complexity thinking to evaluation. *Journal of Environmental Planning and Management*

46:177–197.

Cook, A. and Miller, E. (2012) Talking Points. *Personal Outcome Approach. Practical Guide.* Edinburgh: Joint Improvement Team.

Coulson, A. (2009) Targets and terror: Government by performance indicators. *Local Government Studies*, 35(2), pp. 271-281.

Creswell, J.W. (2007) Five qualitative approaches to inquiry. *Qualitative inquiry and research design: Choosing among five approaches*, 2, pp. 53-80.

Cullen, M.J. (1975) *The statistical movement in early Victorian Britain: the foundations of empirical social research*. Brighton: Harvester Press.

Curry, L.A., Nembhard, I.M. and Bradley, E.H. (2009) Qualitative and mixed methods provide unique contributions to outcomes research. *Circulation*, 119(10), pp. 1442-1452.

Curtis, S. and Riva, M. (2010) Health geographies I: complexity theory and human health. *Progress in Human Geography*, 34(2), pp. 215-223.

Cyert, R.M. and March, J.G., 1963. A behavioral theory of the firm. New Jersey: Englewood Cliffs.

Danermark, B., Ekström, M., Jakobsen, L. and Karlsson, J.C. (2002) *Explaining Society: An introduction to critical realism in the social sciences.* London: Routledge, pp. 73-114.

Darzi, A. (2008) *High quality care for all - NHS Next Stage Review Final Report*. London: Department of Health.

Deming, W.E. (1986) *Out of the Crisis*. Cambridge: MIT Press.

Dennard, L., Richardson, K. and Morçöl, G. (Eds.) (2008) *Complexity and policy analysis: Tools and concepts for designing robust policies in a complex world.* Goodyear: ISCE Publishing

Dent, E.B. (1999) Complexity science: A worldview shift. *Emergence*, 1(4), pp. 5-19.

Department of Health and Social Security (1980) *Inequalities in Health: Report of a Working Group.* London: HMSO.

Design Commission (2014) Restarting Britain: Design and Public Services. *Annual Review of Policy Design*, 2(1), pp. 1-10.

Dewey, J. (1910, reprinted 1991). *How We Think.* New York: Prometheus books

Dewey, J. (1925, reprinted 1998) *The Development of American Pragmatism.* In The Later Works, vol. Boydston Carbondale: Southern Illinois University.

Dewey, J. (1925) *The Development of American Pragmatism*, reprinted in The Essential

Dewey, Volume 1, Pragmatism, Education, Democracy (1998), Edited by Larry A. Hickman and Thomas M. Alexander, Indiana University Press, Bloomington, Indiana, p.3.

Diez Roux, A.V. (2011) Complex systems thinking and current impasses in health disparities research. *American journal of public health*, 101(9), pp. 1627-1634.

Disley, E., Rubin, J., Scraggs, E., Burrowes, N., Culley, D. and Rand, E. (2011) Lessons learned from the planning and early implementation of the Social Impact Bond at HMP Peterborough. *Research series*, 5(11).

Dixon-Woods, M. (2010) Why is patient safety so hard? A selective review of ethnographic studies. *Journal of Health Services Research & Policy*, 15(suppl 1), pp. 11-16.

Dooley, K. (1996) Complex adaptive systems: A nominal definition. *The Chaos Network*, 8(1), pp. 2-3.

Dolan, S.L., García, S. and Auerbach, A. (2003) Understanding and managing chaos in organisations. *International Journal of Management*, 20(1), p.23.

Dopson, S. and Fitzgerald, L. (2005) *Knowledge to action? Evidence-based health care in context*. Oxford University Press.

Dozois, E., Blanchet-Cohen, N. and Langlois, M. (2010) *DE 201: A practitioner's guide to developmental evaluation*. JW McConnell Family Foundation.

Drucker, P.F. (1954) *The Practice Of Management*. Harper & Row. New York.

Dubb, S., McKinley, S. and Howard, T. (2013) *Achieving the Anchor Promise: Improving Outcomes for Low-Income Children, Families and Communities*. Takoma Park: The Democracy Collaborative.

Dunleavy, P. and Hood, C. (1994) From old public administration to new public management. *Public money & management*, 14(3), pp. 9-16.

Dunleavy, P. (2003) *Authoring a PhD: How to plan, draft, write and finish a doctoral thesis or dissertation*. Palgrave Macmillan.

Dunsire, A. (1978) *Control in a Bureaucracy*. Oxford: Martin Robertson.

Dunsire, A. (1996) Tipping the balance: autopoiesis and governance. *Administration & Society*, 28(3), pp. 299-334.

Dunsire, A. (1991) 16 A Cybernetic View of Guidance, Control. *The Public Sector: Challenge for Coordination and Learning*, 31, p.325.

Durkheim, E. (1897, reprinted 1951) *Suicide*. Reprint, Illinois: Free Press.

Durose, C., Needham, C., Mangan, C. and Rees, J. (2017) Generating'good enough'evidence for co-production. *Evidence & Policy: A Journal of Research, Debate and Practice*, 13(1), pp. 135-151.

Easton, G. (2010) Critical realism in case study research. *Industrial marketing management*, 39(1), pp. 118-128.

Easterby-Smith, M., Crossan, M. and Nicolini, D. (2000) Organizational learning: debates past, present and future. *Journal of management studies*, 37(6), pp. 783-796.

Elmore, R. F. (1979) Backward mapping: Implementation research and policy decisions. *Political science quarterly*, *94*(4), pp.601-616.

Elvidge, J. (2011) *The experiences of Scotland's devolved authority can provide lessons in governance for Westminster and beyond*. British Politics and Policy at LSE.

Elvidge, J. (2012) *The Enabling State: A Discussion Paper.* Carnegie UK Trust.

Englebart, D.C. (2003) September*. Improving our ability to improve: A call for investment in a new future.* IBM Co-Evolution Symposium.

Fabricius, C., and G. Cundill. (2014) Learning in adaptive management: insights from published practice. *Ecology and Society* 19(1):29*.*

Ferlie, E.B. and Shortell, S.M., 2001. Improving the quality of health care in the United Kingdom and the United States: a framework for change. *The Milbank Quarterly*, *79*(2), pp.281-315.

Ferlie, E., Fitzgerald, L., McGivern, G., Dopson, S. and Bennett, C. (2011) Public policy networks and 'wicked problems': a nascent solution? *Public Administration*, 89(2), pp. 307-324.

Finegood, D.T., Merth, T.D. and Rutter, H. (2010) Implications of the foresight obesity system map for solutions to childhood obesity. *Obesity*, 18(1), pp. 13-16.

Finegood D.T. (2011) *The Complex Systems Science of Obesity.* In: Cawley J, editor. Handbook of the Social Science of Obesity. Oxford: Oxford University Press

Fink, D.S., Keyes, K.M. and Cerdá, M. (2016) Social Determinants of Population Health: A Systems Sciences Approach. *Current epidemiology reports*, 3(1), pp. 98-105.

Fiol, C.M. and Lyles, M.A. (1985) Organizational learning. *Academy of management review*, 10(4), pp. 803-813.

Fisher, M., Milos, D., Baum, F. and Friel, S. (2016) Social determinants in an Australian urban region: a 'complexity' lens. *Health promotion international*, 31(1), pp. 163-174.

Fitoussi, J.P., Sen, A. and Stiglitz, J.E. (2009) Report by the commission on the measurement of economic performance and social progress. Paris: Government of France.

Fitz-Gibbon, C. T. (1996) *Monitoring education.* London: A&C Black.

Fleetwood, S. and Ackroyd, S. (2004) *Critical realist applications in organisation and management studies* (Vol. 11). Hove: Psychology Press.

Flood, R.L. and Romm, N.R. (1995) Enhancing the process of methodology choice in total systems intervention (TSI) and improving chances of tackling coercion. *Systemic Practice and Action Research*, 8(4), pp. 377-408.

Flyvbjerg, B. (2006) Five misunderstandings about case-study research. *Qualitative inquiry*, *12*(2), pp. 219-245.

Frederickson, H.G., Smith, K.B., Larimer, C.W. and Licari, M. (2015) *The public administration theory primer*. Boulder: Westview Press.

Friedman, M. (2005) *Trying hard is not good enough: How to produce measurable improvements for customers and communities.* Victoria: Trafford.

Gamble, J.A. (2008) *A developmental evaluation primer*. Montreal: JW McConnell Family Foundation.

Gatrell, A.C. (2005) Complexity theory and geographies of health: a critical assessment. *Social Science & Medicine*, 60(12), pp. 2661-2671.

Gell-Mann, M. (1994) *Complex adaptive systems.* In G. Cowan & D. Pines & D. Meltzer (Eds.), Complexity: Metaphors, models and reality (pp. 17–45). Reading: MA: Addison-Wesley Publishing Company.

Gell-Mann, M. (1995) *The Quark and the Jaguar: Adventures in the Simple and the Complex*. Macmillan.

Gerrits, L. and Verweij, S. (2013) Critical realism as a meta-framework for understanding the relationships between complexity and qualitative comparative analysis. *Journal of Critical Realism*, 12(2), pp. 166-182.

Gershenson, C. (2015) Requisite variety, autopoiesis, and self-organization, *Kybernetes*, Vol. 44 Issue: 6/7, pp. 866-873, doi: 10.1108/K-01-2015-0001

Geyer, R. and Cairney, P. eds. (2015) *Handbook on complexity and public policy*. Cheltenham: Edward Elgar Publishing.

Geyer, R.R. and Cairney, P. (2016) *A critical discussion of complexity theory: how does' complexity thinking'improve our understanding of politics and policymaking?* (Forthcoming) Complexity, Governance and Networks.

Ghandour, R.M., Flaherty, K., Hirai, A., Lee, V., Walker, D.K. and Lu, M.C. (2017) Applying Collaborative Learning and Quality Improvement to Public Health: Lessons from the Collaborative Improvement and Innovation Network (CoIIN) to Reduce Infant Mortality. *Maternal and Child Health Journal*, pp. 1-9.

Giddens, A. (1984) *The constitution of society: Outline of the theory of structuration.* Cambridge: Polity Press.

Gilson, C., Dunleavy, P. and Tinkler, J. (2009) *Organizational Learning in Government Sector Organizations: Literature Review.* LSE Public Policy Group. London.

Gilstrap, D.L. (2005) Strange attractors and human interaction: Leading complex organizations through the use of metaphors Complicity: *An International Journal of Complexity and Education Vol.* 2, no. 1: 55-69.

Gigerenzer, G. (2015) *Risk savvy: How to make good decisions*. London: Penguin.

Glansdorff, P. and Prigogine, I. (1971) *Structure, stability and fluctuations.* New York: Interscience.

Glaser, B. and Strauss, A. (1967). *The discovery of grounded theory*. *London: Weidenfield &* Nicolson.

Glendinning, C., Arksey, H., Jones, K., Moran, N., Netten, A. and Rabiee, P. (2009) *The individual budgets pilot projects: impact and outcomes for carers.* York: Social Policy Research Unit.

Goddard, J. (2009) *Re-inventing the Civic University*. London: Nesta.

Godwin, W. (1820) *Of Population: an Enquiry concerning the power of increase in the numbers of mankind, being an answer to Mr. Malthus's Essay on that subject*. London: Longman, Hurst, Rees, Orme & Brown.

Golafshani, N. (2003) Understanding reliability and validity in qualitative research. *The qualitative report*, 8(4), pp. 597-606.

Greve, H.R. (2003) A behavioral theory of R&D expenditures and innovations: Evidence from shipbuilding. *Academy of management journal*, *46*(6), pp. 685-702.

Gruber, M., De Leon, N., George, G. and Thompson, P. (2015) Managing by design. *Academy of Management Journal*, 58(1), pp. 1-7.

Guba, E.G. (1981) Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational technology research and development*, 29(2), pp. 75-91.

Gubrium, J. F., & Holstein, J. A. (2002) From the individual interview to the interview society. In J. F. Gubrium & J. A. Holstein (Eds.), *Handbook of interview research: Context and method* (pp. 3–32).Thousand Oaks: Sage.

Hambleton, R. and Howard, J. (2013) Place-based leadership and public service innovation. *Local Government Studies*, 39(1), pp. 47-70.

Hanf, K. I. and Scharpf, F. W. (Eds.) (1978). *Interorganizational policy making*. London: Sage.

Hargreaves, A. and Moore, S. (2000) Educational outcomes, modern and postmodern interpretations: Response to Smyth and Dow. *British Journal of Sociology of Education*, 21(1), pp. 27-42.

Harvey, D.L. and Reed, M. (1996) Social science as the study of complex systems. *Chaos theory in the social sciences*, pp. 295-324.

Hawk, T.F. and Zand, D.E. (2014) Parallel Organization: Policy Formulation, Learning, and Interdivision Integration. *The Journal of Applied Behavioral Science*, 50(3), pp. 307-336.

Haynes, P. (2015) *Managing complexity in the public services*. London: Routledge.

Head, B.W. and Alford, J. (2015) Wicked problems: Implications for public policy and management. *Administration & Society*, *47*(6), pp. 711-739.

Heifetz, R.A. and Laurie, D.L. (1997) The work of leadership. *Harvard Business Review*, *75*, pp. 124-134.

Heinrich, C.J. (2002) Outcomes–based performance management in the public sector: implications for government accountability and effectiveness. *Public administration review*, *62*(6), pp. 712-725.

Hirschman, A.O. (1970) *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states* (Vol. 25). Cambridge: Harvard university press.

HM Government. (2011a) *Open public Services White Paper*. London: The Stationary Office.

HM Treasury (2010) *Total Place: a whole area approach to public services*. March. London: HM Treasury.

Holland, J.H. (1992) Complex adaptive systems. *Daedalus*, pp. 17-30.

Holland, J.H. (1995) *Hidden order: How adaptation builds complexity*. New York: Basic Books.

Holland, J.H. (2006) Studying complex adaptive systems. *Journal of systems science and complexity*, 19(1), pp. 1-8.

Hood, C. (1976) *The limits of administration*. London: Wiley.

Hood, C. (1991) A public management for all seasons?. *Public administration*, 69(1), pp. 3-19.

Hood, C. (1995) The "New Public Management" in the 1980s: variations on a theme. *Accounting, organizations and society*, 20(2-3), pp. 93-109.

Hood, C. (2007) Public service management by numbers: Why does it vary? Where has it come from? What are the gaps and the puzzles?. *Public Money and Management*, 27(2), pp. 95-102.

Hood, C. (2010) Can we? Administrative limits revisited. *Public Administration Review*, 70(4), pp. 527-534.

Hood, C., James, O., Peters, B.G. and Scott, C. (2004) *Controlling modern government: Variety, commonality and change* (pp. 3-21). Cheltenham.

Hood, C. and Dixon, R. (2016) Not What It Said on the Tin? Reflections on Three Decades of UK Public Management Reform. *Financial Accountability & Management*, 32(4), pp. 409-428.

Hood, R. (2012) A critical realist model of complexity for interprofessional working. *Journal of interprofessional care*, 26(1), pp. 6-12.

Housden, P. (2014) This is us: A perspective on public services in Scotland. *Public Policy and Administration*, 29(1), pp. 64-74.

Housden, P. (2016) *Rethinking Public Services.* Centre for Public Impact. Available from: https://s3-eu-west-1.amazonaws.com/funct-cpi-wordpress/assets/uploads/2016/07/Rethinking-Public-Services-A4_Final.pdf [accessed 25 April 2017].

Hulme, D. (2010) Lessons from the making of the MDGS: human development meets results-based management in an unfair world. *IDS bulletin*, 41(1), pp. 15-25.

Ilic, M. and Bediako, S. (2011). Project Oracle. In R. Puttick (Ed.), *Evidence for Social Policy and Practice: Perspectives on how research and evidence can influence decision making in public services* (pp. 4-9). London: NESTA.

Inkelas, M. and Bowie, P. (2014) The Magnolia Community Initiative: The Importance of Measurement in Improving Community Well-Being. *Community Investments*, (01), pp. 18-24.

Inkelas, M. and McPherson, M.E. (2015) December. Quality improvement in population health systems. In *Healthcare* (Vol. 3, No. 4, pp. 231-234). Amsterdam: Elsevier.

Innes, J. E., & Booher, D. E. (1999) Consensus building and complex adaptive systems: A framework for evaluating collaborative planning. Journal of the American Planning Association, 65(4), 412-423.

Institute for Healthcare Improvement (2003) *The Breakthrough Series: IHI's Collaborative Model for Achieving Breakthrough Improvement. IHI Innovation Series white paper*. Boston: Institute for Healthcare Improvement

IRISS (2015) *The view from here.* Glasgow: IRISS.

Jaaron, A.A. and Backhouse, C.J. (2016) Operationalising "Double-Loop" Learning in Service Organisations: A Systems Approach for Creating Knowledge. *Systemic Practice and Action Research*, pp. 1-21.

Jackson, P.M. (2011) Governance by numbers: what have we learned over the past 30 years?. *Public Money & Management*, 31(1), pp. 13-26.

James, S. (1993) The idea brokers: the impact of think tanks on British government. *Public administration*, 71(4), pp. 491-506.

Jantsch, E. (1980) *The Self-Organizing Universe*. Oxford: Pergamon Press.

Jayasinghe, S. (2011) Conceptualising population health: from mechanistic thinking to complexity science. *Emerging Themes in Epidemiology*, 8(1), p.2.

Jayasinghe, S. (2015) Social determinants of health inequalities: towards a theoretical perspective using systems science. *International journal for equity in health*, *14*(1), p.71.

Jensen, K. B. (2002). *The Qualitative Research Process'. In: K. B. Jensen (ed.), A Handbook of Media and Communication Research – Qualitative and Quantitative Methodologies*, pp. 235–254. London: Routledge

Jessop, B. (1997) The governance of complexity and the complexity of governance: preliminary remarks on some problems and limits of economic guidance. In: *Beyond*

market and hierarchy: interactive governance and social complexity, pp. 95-128. Cheltenham: Edward Elgar Publishing.

Jessop, B. (1998) *The rise of governance and the risks of failure: the case of economic development. International social science journal,* 50(155), pp. 29-45.

Jick, T.D. (1979) Mixing qualitative and quantitative methods: Triangulation in action. *Administrative science quarterly*, 24(4), pp. 602-611.

Jones, G. and Stewart, J. (2009) New development: Accountability in public partnerships—The case of Local Strategic Partnerships. *Public Money & Management*, 29(1), pp. 59-64.

Kennedy, M.M. (1979) Generalizing from single case studies. *Evaluation quarterly*, 3(4), pp. 661-678.

Kickert, W. J. M., Klijn, E. H. and Koppenjan, J. F. M. (Eds.) (1997) *Managing complex networks: Strategies for the public sector*. London: SAGE.

Kiel, L. D. (1989) *Nonequilibrium theory and its implications for public administration*. Public Administration Review, 49(6), 544–551.

Kiel, L. D. (1993) *Nonlinear dynamical analysis: Assessing systems concepts in a government agency.* Public Administration Review, 53(2), 143–153.

Kilmann, R.H. (1982) Designing collateral organizations. *Human Systems Management*, 3(2), pp. 66-76.

Kilo, C.M. (1998) A framework for collaborative improvement: lessons from the Institute for Healthcare Improvement's Breakthrough Series. *Quality Management in Healthcare*, 6(4), pp. 1-14.

Kindig, D. and Stoddart, G. (2003) What is population health?. *American journal of public health*, 93(3), pp. 380-383.

King, C. and Jiggins, J. (2002) A systemic model and theory for facilitating social learning. *Wheelbarrows full of frogs. Social learning in rural resource management*, pp. 85-104.

Kitson, A.L. (2009) The need for systems change: reflections on knowledge translation and organizational change. *Journal of advanced nursing*, 65(1), pp. 217-228.

Klijn, E. H. (2008b) Complexity theory and public administration: What's new? Key concepts in complexity theory compared to their counterparts in public administration research. *Public Management Review*, 10(3), 299-317.

Klijn, E. H., & Koppenjan, J. (2012). Governance network theory: past, present and future. *Policy & Politics*, 40(4), 587-606.

Koch, U., Stout, S., Landon, B.E. and Phillips, R.S. (2016) From healthcare to health: A proposed pathway to population health. *In Healthcare* (Vol. 4, No. 4, pp. 291-297). Elsevier.

Kolb, D. A. (1984) *Experiential Learning. Experience as the source of learning and development.* Englewood Cliffs: Prentice-Hall.

Krumholz, H.M. (2008) Outcomes research. *Circulation*, 118(3), pp. 309-318.

Kurtz, C. F. and Snowden, D. J. (2003) *The new dynamics of strategy: Sense-making in a complex and complicated world.* IBM *Systems Journal*, Volume 42, Number 3, 462

Kvale, S. (1996) Interviews*: an introduction to qualitative research interviewing*. London: Sage.

Laegreid, P. and Christensen, T. eds. (2013.) *Transcending new public management: the transformation of public sector reforms*. Farham: Ashgate Publishing.

Langley, G.J., Moen, R.D., Nolan, K.M., Nolan, T.W., Norman, C.L. and Provost, L.P. (2009) *The improvement guide: a practical approach to enhancing organizational performance*. John Wiley & Sons.

Lassey, P. (1998). *Developing a Learning Organization*. London: Kogan Page.

Law, J. (2013). *Do Outcomes Based Approaches to Service Delivery Work?* Local Authority Outcome Agreements in Wales. Cardiff: University of South Wales.

Le Grand, J. (1991) Quasi-markets and social policy. *The Economic Journal*, 101(408), pp. 1256-1267.

Le Grand, J. (2003) *Motivation, agency, and public policy: of knights and knaves, pawns and queens*. Oxford: Oxford University Press on Demand.

Le Grand, J. (2009) *The other invisible hand: Delivering public services through choice and competition*. Princeton: Princeton University Press.

Lecy, J.D., Mergel, I.A. and Schmitz, H.P. (2014) Networks in public administration: current scholarship in review. *Public Management Review*, 16(5), pp. 643-665.

LeMahieu, P.G., Grunow, A., Baker, L., Nordstrum, L.E. and Gomez, L.M. (2017) Networked Improvement Communities: the discipline of improvement science meets the power of networks. *Quality Assurance in Education*, 25(1).

Lewis, J. and Ritchie, J. (2003) Generalising from qualitative research. In: *Qualitative research practice: A guide for social science students and researchers*, pp. 263-286. London: Sage.

Lichtenstein, B.B. and Plowman, D.A. (2009) The leadership of emergence: A complex systems leadership theory of emergence at successive organizational levels. *The Leadership Quarterly*, 20(4), pp. 617-630.

Lincoln, Y. S. and Guba, E. G. (1985) *Naturalistic inquiry*. Beverly Hills: Sage.

Lindblom, C. E. (1959). *The science of" muddling through". Public administration review,* 79-88.

Ling, T. (2002) Delivering joined–up government in the UK: dimensions, issues and problems. *Public administration*, 80(4), pp. 615-642.

Loeffler, E., Power, G., Bovaird, T.,Hine-Hughes, F., eds (2013) *Co-production of health and wellbeing in Scotland*. Birmingham: Governance International.

Loeffler, E. and Bovaird, T. (2016) User and Community Co-Production of Public Services: What Does the Evidence Tell Us? *International Journal of Public Administration*, 39(13), pp. 1006-1019.

Lowe, T. (2013) New development: The paradox of outcomes—the more we measure, the less we understand. *Public Money & Management*, 33(3), pp. 213-216.

Lowe, T. *Debate: Complexity and the performance of social interventions.* Public Money & Management, 37(2), p.79-80.

Lowe, T. and Wilson, R. (2015) Playing the Game of Outcomes-based Performance Management. Is Gamesmanship Inevitable? Evidence from Theory and Practice. *Social Policy & Administration.*

Lowe, T. and Wilson, R. (2016) The Performance Management of complex systems – enabling adaptation. Conference paper presented at the International Research Society for Public Management conference, June 2016.

Lowe, T., Wilson, R. and Boobis, S. 2016. The performance management of complex systems – enabling adaptation. Conference paper presented at the Performance Management Association conference, April 2016.

Lowndes, V. and McCaughie, K. (2013) Weathering the perfect storm? Austerity and institutional resilience in local government. *Policy & Politics*, 41(4), pp. 533-549.

Luhmann, N. (1995) *Social systems*. Stanford University Press.

Lusch, R.F. and Vargo, S.L. (2006) Service-dominant logic: reactions, reflections and refinements. *Marketing theory*, 6(3), pp. 281-288.

Mair, C., Zdeb, K. and Markie, K. (2015) *Making better places: making places better – the distribution of positive and negative outcomes in Scotland.* Broxburn: Improvement Service.

Mair, C. (2016) The "outcome focus": a Scottish reflection. Carnegie blog. Available: http://www.carnegieuktrust.org.uk/blog/outcome-focus-scottish-reflection/.

March, J.G. (1991) Exploration and exploitation in organizational learning. *Organization science*, 2(1), pp. 71-87.

Marmot, M. (2005) Social determinants of health inequalities. *The Lancet*, 365(9464), pp. 1099-1104.

Marmot, M., Friel, S., Bell, R., Houweling, T.A., Taylor, S. and Commission on Social Determinants of Health. (2008) Closing the gap in a generation: health equity through action on the social determinants of health. *The Lancet*, 372(9650), pp. 1661-1669.

Mason, J. (2002) *Qualitative researching (2nd ed.).* London: Sage.

Mayne, J. (2001) Addressing attribution through contribution analysis: using performance measures sensibly. *The Canadian journal of program evaluation*, 16(1), p.1.

Mayne, J. (2007) Challenges and lessons in implementing results-based management. *Evaluation*, 13(1), pp. 87-109.

Maxwell, J. A. (2005) *Qualitative research design: An interactive approach: Vol. 41. Applied social science research methods series (2nd ed.).* Thousand Oaks, CA: Sage.

Maxwell, J.A. (2008) The value of a realist understanding of causality for qualitative research. *Qualitative inquiry and the politics of evidence*, pp. 163-181.

Maxwell, J.A. (2015) A Critical Realist Perspective for Qualitative Research. *Qualitative Inquiry—Past, Present, and Future: A Critical Reader*, 88.

McAteer, M. and Stephens, A. (2013) New development: The role of benchmarking in supporting improvement in local government—Scottish and Welsh practitioners' perspectives. *Public Money & Management*, 33(4), pp. 281-284.

McElroy, M. (2000) Integrating complexity theory, knowledge management and organizational learning. *Journal of Knowledge Management*, 4 (3),195–203.

McHugh, N., Sinclair, S., Roy, M., Huckfield, L. and Donaldson, C. (2013) Social impact bonds: a wolf in sheep's clothing?. *Journal of Poverty and Social Justice*, 21(3), pp. 247-257.

McPherson, M.E., Gloor, P.A. and Smith, L.A. (2015) Using collaborative improvement and innovation networks to tackle complex population health problems. *JAMA pediatrics*, 169(8), pp. 709-710.

Meadows, D. (1999) *Leverage points. Place to intervene in a system.*
Hartland: The Sustainability Institute.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass.

Merry, U. (1999) Organizational strategy on different landscapes: a new science approach. In: *Systemic Practice and Action Research*, vol. 12, no. 3, p. 257-278.

Midwinter, A. (2009) Scotland's concordat: an assessment of the new financial framework in central–local relations. Public Money & Management, 29, 1, pp. 65–70.

Miles, M.B. and Huberman, A.M. (1994) *Qualitative data analysis: A sourcebook*. Beverly Hills: Sage Publications.

Mill, J.S. (1884) *A system of logic ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation*. Harper.

Miller, E. (2010) Can the shift from needs-led to outcomes-focused assessment in health and social care deliver on policy priorities?. *Research, Policy and Planning*, 28(2), pp. 115-127.

Miller, E. (2014) Considering two outcomes paradigms: the improving (person-centred) and the proving (managerialist) agendas. In: *Human Rights and Social Equality: Challenges for Social Work.* Ashgate Publishing. pp. 34-39

Mingers, J. (2011) `The Contribution of Systematic Thought to Critical Realism, *Journal of Critical Realism*, Vol.10, No. 3, 303-330.

Mintzberg, H. (1994) The fall and rise of strategic planning. *Harvard business review*, 72(1), pp. 107-114.

Moore, M.H. (1995) *Creating public value: Strategic management in government*. Cambridge: Harvard University Press.

Morçöl, G. (2013) *A complexity theory for public policy*. London: Routledge.

Morgan, Gareth. (1997) *Images of organization (2nd ed.)*. Thousand Oaks, CA: Sage Publications.

Morse, J.M., Barrett, M., Mayan, M., Olson, K. and Spiers, J. (2002) Verification strategies for establishing reliability and validity in qualitative research. *International journal of qualitative methods*, 1(2), pp. 13-22.

Morse, A. (2015) *Outcome-based payment schemes: Government's use of payment by results* (Vol. 6/15). London: National Audit Office.

Moynihan, D.P. (2005) Goal-based learning and the future of performance management. *Public Administration Review*, 65(2), pp. 203-216.

Mulhall, A. (2003) In the field: notes on observation in qualitative research. *Journal of advanced nursing*, 41(3), pp. 306-313.

Murphy, J., Rhodes, M.L., Meek, J.W. and Denyer, D. (2016) *Managing the Entanglement: Complexity Leadership in Public Sector Systems*. *Public Administration Review.*

National Collaborating Centre for Determinants of Health (2007). *Making Sense of It All*

Conducting KSTE with Canadian Public Health. Nova Scotia: National Collaborating Centre

for Determinants of Health.

Needham, C. and Carr, S. (2009) *Co-production-an Emerging Evidence Base for Adult Social Care Transformation: Research Briefing*. Social Care Institute for Excellence.

Nicholas, E., Qureshi, H. and Bamford, C. (2003) *Outcomes into Practice: Focusing Practice and Information on the Outcomes People Value*. A Resource Pack for Managers and Trainers. York: Social Policy Research Unit.

Niskanen, W.A. (1975) Bureaucrats and politicians. *The Journal of Law and Economics*, 18(3), pp. 617-643.

Nocon, A. and Qureshi, H. (1996) *Outcomes of community care for users and carers: a social services perspective*. Open Univ Pr.

Normann, R. (1984), *Service Management.* New York: John Wiley& Sons.

Northern Ireland Assembly. (2016) *Research and Information Service Briefing Paper: Outcomes-based government.* Available from:

http://www.niassembly.gov.uk/globalassets/documents/raise/publications/2016-2021/2016/executive_office/4116.pdf

Nutley, S., Powell, A. and Davies, H. (2013) *What counts as good evidence? provocation paper for the alliance for useful evidence*. London: Alliance for Useful Evidence.

Ogrinc G, Shojania KG. (2014) Building knowledge, asking questions. BMJ Qual Saf 23(4):265-7.

Organisation for Economic Co-operation and Development. (2007) *Quality and Equity of Schooling in Scotland*. Paris: OECD

Organisation for Economic Co-operation and Development (2011) *Together for Better Public Services: Partnering with Citizens and Civil Society*. Paris: OECD.

Organisation for Economic Co-operation and Development (2017) From Transactional to Strategic: Systems Approaches to Public Challenges. Paris: OECD

Olson, M. (1965) *The logic of collective action.* Cambridge: Mass. Harvard.

Olsson, P., C. Folke, and F. Berkes (2004) Adaptive comanagement for building resilience in

social–ecological systems. *Environmental Management* (34), pp. 75–80.

Osborne, D. and Gaebler, T. (1992) *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector.* New York: Penguin Books

Osborne, S. (2006) Editorial: The new public governance? *Public Management Review,* 8 (3), 377-387.

Osborne, S. P., Radnor, Z. and Nasi, G. (2013) A new theory for public service management? Toward a (public) service-dominant approach. *The American Review of Public Administration*, 43(2), 135-158.

Osborne, S.P. and Strokosch, K. (2013) It takes two to tango? understanding the Co-production of public services by integrating the services management and public administration perspectives. *British Journal of Management*, 24(S1), pp. S31-S47.

Osborne, S.P. and Radnor, Z. (2016) The New Public Governance and Innovation in Public Services. *Enhancing Public Innovation by Transforming Public Governance*, p.54.

Osborne, S.P., Radnor, Z. and Strokosch, K. (2016) Co-Production and the Co-Creation of Value in Public Services: A suitable case for treatment?. *Public Management Review*, 18(5), pp. 639-653.

Osborne, S.P. (2017) Public management research over the decades: what are we writing about?

Ostrom, V. and Ostrom, E. (1971) Public choice: A different approach to the study of public administration. *Public Administration Review*, 31(2), pp. 203-216.

Ostrom, E., Parks, R.B., Whitaker, G.P. and Percy, S.L. (1978) The public service production process: a framework for analyzing police services. *Policy Studies Journal*, 7(s1), pp. 381-381.

Ostrom, E. (2010) Beyond markets and states: polycentric governance of complex economic systems. *Transnational Corporations Review*, 2(2), pp. 1-12.

Øvretveit, J., Bate, P., Cleary, P., Cretin, S., Gustafson, D., McInnes, K., McLeod, H., Molfenter, T., Plsek, P., Robert, G. and Shortell, S., 2002. Quality collaboratives: lessons from research. *Quality and safety in health care*, *11*(4), pp. 345-351.

Paley, C. and Slasberg, C. (2007) Implementing outcome-based commissioning. *Journal of Care Services Management*, 1(4), pp. 353-361.

Palmberg, K. (2009) Complex adaptive systems as metaphors for organizational management. *The Learning Organization*, 16(6), pp. 483-498.

Parks, R.B., Baker, P.C., Kiser, L., Oakerson, R., Ostrom, E., Ostrom, V., Percy, S.L., Vandivort, M.B., Whitaker, G.P. and Wilson, R. (1981) Consumers as coproducers of public services: Some economic and institutional considerations. *Policy Studies Journal*, 9(7), pp. 1001-1011.

Pascale, Richard T., Millemann, M. and Gioja, L. (2000) *Surfing the edge of chaos: The laws of nature and the new laws of business.* New York: Crown Business.

Parker, D., Stacey, R. D. and Robinson, C. (1994) *Chaos, management and economics: The implications of non-linear thinking* (Vol. 125). London: Institute of Economic Affairs.

Parsons, T. (1949) *The structure of social action* (Vol. 491). New York: Free Press.

Patton, M.Q., 1994. Developmental evaluation. *Evaluation practice*, *15*(3), pp. 311-319.

Patton, M. Q. (2002) *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks,CA: Sage Publications, Inc.

Pawson, R. and Tilley, N. (1997) *Realistic evaluation*. London: Sage.

Pearce, N. and Merletti, F. (2006) Complexity, simplicity, and epidemiology.

Pedder, D. and MacBeath, J. (2008) Organisational learning approaches to school leadership and management: teachers' values and perceptions of practice. *School Effectiveness and School Improvement*, 19(2), pp. 207-224.

Peirce, C.S. (1957) The logic of abduction. In: *Peirce's essays in the philosophy of science*. New York: The Liberal Science Press

Perrin, B. (2006) *Moving from outputs to outcomes: Practical advice from governments around the world*. IBM Center for the Business of Government.

Petch, A., Cook, A. and Miller, E. (2013) Partnership working and outcomes: do health and social care partnerships deliver for users and carers?. *Health & social care in the community*, 21(6), pp. 623-633.

Plsek, P.E. (1999) Quality improvement methods in clinical medicine. *Pediatrics*, *103*(Supplement E1), pp. 203-214.

Polit, D.F., Beck, C.T. and Hungler, B.P. (2001). *Essentials of Nursing Research: Methods, Appraisal, and Utilisation* (5th edn). Philadelphia: Lippincott.

Pollitt, C. (1996) Business approaches to quality improvement: why they are hard for the NHS to swallow. *Quality in health care, 5(*2), p.104.

Pollitt, C. and Bouckaert, G. (2004) *Public management reform: A comparative analysis*. Oxford: Oxford University Press.

Pollitt, C. (2009). Complexity theory and evolutionary public administration: A skeptical afterword. In G. Teisman, A. Van Buuren, & L. Gerrits (Eds.), *Managing complex governance systems*. London: Routledge.

Pollitt, C. (2013) The evolving narratives of public management reform: 40 years of reform white papers in the UK. *Public Management Review*, 15(6), pp. 899-922.

Pollitt, C. (2015) Wickedness will not wait: climate change and public management research. *Public Money & Management*, 35(3), pp. 181-186.

Pollitt, C. (2016) Losing the Big Picture, and Influence Too. Statecrafting Blog. Available: *https://statecrafting. net/losing-the-big-picture-and-losing-influence-too-f4a2d8e57532*.

Popper, K.R. (1969) *Conjectures and Refutations: The Growth of Scientific Knowledge* (3rd edn.). Routledge & Kegan Paul: London.

Porter, T.M. (1986) *The rise of statistical thinking, 1820-1900*. Princeton: Princeton University Press.

Prigogine, I. (1967) *Introduction to thermodynamics of irreversible processes*. New York: Interscience*.*

Prigogine, I., Stengers, I. and Prigogine, I. (1984) *Order out of chaos: Man's new dialogue with nature* (Vol. 13). New York: Bantam books.

Putnam, H. (1981) *"Two Philosophical Perspectives," chap. 3 of Reason, Truth and History*. Cambridge: Cambridge University Press.

Puttick, R. and Ludlow, J. (2013) *Standards of Evidence: An approach that balances the need for evidence with innovation*. London: Nesta.

Qureshi, H., Patmore, C., Nicholas, E. and Bamford, C. (1998) Overview: outcomes of social care for older people and carers. *Outcomes in Community Care Practice*, 5(5), pp. 12-15.

Qureshi, H. and Nicholas, E. (2001) A new conception of social care outcomes and its practical use in assessment with older people. *Research, Policy and Planning*, pp. 11-26.

Radnor, Z., Osborne, S. P., Kinder, T., & Mutton, J. (2014). Operationalizing Co-Production in Public Services Delivery: The contribution of service blueprinting. *Public Management Review*, 16(3), 402-423.

Raisch, S. and Birkinshaw, J. (2008) Organizational ambidexterity: Antecedents, outcomes, and moderators. *Journal of Management*, 34(3), pp. 375-409.

Reed, M. and Harvey, D.L. (1992) The new science and the old: Complexity and realism in the social sciences. *Journal for the Theory of Social Behaviour*, 22(4), pp. 353-380.

Reed, M., Evely, A.C., Cundill, G., Fazey, I.R.A., Glass, J., Laing, A., Newig, J., Parrish, B., Prell, C., Raymond, C. and Stringer, L. (2010) What is social learning?. *Ecology and Society*.

Reed, J.E. and Card, A.J. (2016) The problem with plan-do-study-act cycles. *BMJ Qual Saf*, 25(3), pp. 147-152.

Rhodes, R.A. (1997) *Understanding governance: Policy networks, governance, reflexivity and accountability.* Milton Keynes: Open University Press*.*

Rhodes, M.L. (2008) Complexity and emergence in public management: The case of urban regeneration in Ireland. *Public Management Review*, 10(3), pp. 361-379.

Rhodes, M. L. and MacKechnie, G. (2003) *Understanding public service systems: Is there a role for complex adaptive systems theory?* Emergence, 5(4), 57-85.

Rhodes, M.L., Murphy, J., Muir, J. and Murray, J.A. (2010) *Public management and complexity theory: Richer decision-making in public services*. Routledge.

Rhodes, M.L. (2013) Innovation in complex public service systems. *Handbook of Innovation in Public Services*, p.332.

Rhodes, M.L., Murphy, J., Muir, J. and Murray, J.A. (2010) *Public management and complexity theory: Richer decision-making in public services*. Routledge.

Richards, L. 2005. *Handling qualitative data: A practical guide.* London: Sage.

Richards, L. and Richards, T. (1994) 'From filing cabinet to computer' in A. Bryman & R.G. Burgess (eds) *Analyzing Qualitative Data,* London: Routledge?

Richardson, K.A. (2008) Managing complex organizations: Complexity thinking and the science and art of management. *Emergence: Complexity and Organization*, 10(2), p.13.

Rihoux, B. and Ragin, C.C. (2009) *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*. Sage.

Ritchie, J., and Lewis, J. (2003) *Qualitative research practice: A guide for social science students and researchers*. London: Sage.

Rittel, H. W., & Webber, M. M. (1973) Dilemmas in a general theory of planning. *Policy sciences*, 4(2), 155-169.

Roberts, N. (2000) Wicked problems and network approaches to resolution. *International public management review*, 1(1), pp. 1-19.

Roberts, A.S. (2014) *Large Forces: What's Missing in Public Administration*. North Charleston: CreateSpace Independent Publishing Platform.

Rogers, P.J. (2008) Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14(1), pp. 29-48.

Rose, G. (2001) Sick individuals and sick populations. *International journal of epidemiology*, 30(3), pp. 427-432.

Rosenhead, J. (2006) Past, present and future of problem structuring methods. *Journal of the operational research society*, 57(7), pp. 759-765.

Rouse, W. B. (2008) *Health care as a complex adaptive system: implications for design and management.* BRIDGE-WASHINGTON-NATIONAL ACADEMY OF ENGINEERING-, 38(1), 17.

Royal Cornwall Hospitals (2015). *Infection Prevention and Control Roles and Responsibilities Policy.* Available: http://www.rcht.nhs.uk/DocumentsLibrary/RoyalCornwallHospitalsTrust/Clinical/InfectionPreventionAndControl/InfectionPreventionAndControlRolesAndResponsibilitiesPolicy.pdf

Ryan, B. (2003) 'Harder yards ahead: the second stage of public sector reform in New Zealand', *International Review of Public Administration*, 8 (1).

Ryan, A. and O'Malley, L. (2016). The role of the boundary spanner in bringing about innovation in cross-sector partnerships. *Scandinavian Journal of Management*, *32*(1), pp.1-9.

Ryle, G. (1949) *Concept of the mind*. London: Hutchinson and Company.

Sandfort, J.R. (2000) Moving beyond discretion and outcomes: Examining public management from the front lines of the welfare system. *Journal of Public Administration Research and Theory: J-PART*, pp. 729-756.

Sayer, A. (1984) *Method in social science*. London: Routledge.

Sayer, A. (2004) *Why critical realism?* (pp. 6-20). in Fleetwood, S. and Ackroyd, S., *Critical realist applications in organisation and management studies* (Vol. 11). Psychology Press.

Schedler, K. and Proeller, I. (2010) *Outcome-oriented public management: a responsibility-based approach to the new public management*. USA: Age Publishing Incorporated.

Schensul J. J. (2009) Community, culture and sustainability in multilevel dynamic systems intervention science. *American Journal of Community Psychology*, (43), 241–256

Schneider, M. and Somers, M. (2006) Organizations as complex adaptive systems: Implications of complexity theory for leadership research. *The Leadership Quarterly,* 17(4), pp. 351-365.

Scholten, P. (2006) *Social Return on Investment: A Guide to SROI Analysis.* Lenthe Publications.

Schumacher, E.F. (1995) *A Guide for the Perplexed*. Random House.

Schwandt, T.A. (1997) *Qualitative inquiry: A dictionary of terms*. Sage Publications.

Scottish Government (2008) *Getting it Right for Every Child.* Edinburgh: *Scottish* Government.

Scottish Government (2007) '*Concordat between the Scottish Government and local government*'. Edinburgh: Scottish Government.

Scottish Government (2008). *The Early Years Framework.* Edinburgh: Scottish Government.

Scottish Government (2012). *The Early Years Taskforce Shared Vision and Priorities*. Edinburgh: Scottish Government.

Scottish Government (2014) *The Early Years Collaborative (EYC), Stock Take Review of Years 1 and 2*, Edinburgh: Scottish Government.

Scottish Government (2013) *Learning Session Two Delegate Pack*. Edinburgh: Scottish Government. Available: http://www.gov.scot/Resource/0041/00413564.pdf

Scottish Parliament (2014) Evidence given to the Health and Sport Committee. May 2014. Available: http://http//www.scottish.parliament.uk/parliamentarybusiness/28862.aspx?r=9151&mode=pdf

Scottish Neighbourhood Statistics (2012) *Scottish Index of Multiple Deprivation* Available: http://www.sns.gov.uk/

Seddon, J. (2003) *Freedom from command and control*. Vanguard Education Limited.

Seddon, J. (2008) *Systems thinking in the public sector: The failure of the reform regime... and a manifesto for a better way*. Triarchy Press Limited.

Sen, Amartya. (2000) A Decade of Human Development. *Journal of Human Development*, no. 1: 17-23.

Senge, P. (1990) *The fifth discipline: The art and science of the learning organization*. New York: Currency Doubleday.

Senge, P. M., Kleiner, A., Roberts, C., Ross, R. B. and Smith, B. J. (1994) *The fifth discipline fieldbook: Strategies and tools for Building a learning organization*. New York: Doubleday.

Shojania, K. G. and Grimshaw, J. M. (2005). Evidence-based quality improvement: the state of the science. Health affairs, 24(1), 138-150.

Simmons, R. (2009) Understanding The 'Differentiated Consumer' in Public Services. In: Simmons R, Powell M, Greener I (ed.). *he consumer in public services: Choice, values and difference* , Bristol: The Policy Press, pp. 57-76.

Simmons, R. and Birchall, J. (2005) A joined-up approach to user participation in public services: strengthening the "Participation Chain". *Social Policy & Administration*, 39(3), pp. 260-283.

Simon, H. A. *(*1957) *Administrative Behaviour (*2nd edn*).* New York*:* Macmillan.

Sinclair, J. (1798) *The Statistical Account of Scotland*, 21 vols. Edinburgh: William Creech.

Sinclair, Address to the Clergy of the Church of England (1797) Appendix F. *Communications to the Board of Agriculture on Subjects Relative to the Husbandry, and Internal Improvement of the County, Vol 1*. London: Bulmer and Co, p. xxxv.

Smith, P. (1995) Performance indicators and outcome in the public sector. *Public Money & Management*, 15(4), pp. 13-16.

Smyth, J. and Dow, A. (1998.) What's wrong with outcomes? Spotter planes, action plans, and steerage of the educational workplace. *British Journal of Sociology of Education*, 19(3), pp. 291-303.

Snell, R. and Chak, A.M.K. (1998) The learning organization: learning and empowerment for

Snowden, D.J. and Boone, M.E. (2007) A leader's framework for decision making. *Harvard business review*, 85(11), p.68.

Social Return on Investment Working Group (2011) *Outcome-Based Government: How to improve spending decisions across government.* London: Centre for Social Justice

Sørensen, E. and Torfing, J. (2011) Enhancing collaborative innovation in the public sector. *Administration & Society*, 43(8), pp. 842-868.

Sosu, E. and Ellis, S. (2014) *Closing the attainment gap in Scottish education*. Joseph Rowntree Foundation, York.

Stacey, R.D. (1995) The science of complexity: An alternative perspective for strategic change processes. *Strategic management journal*, 16(6), pp. 477-495.

Stacey, R.D. (1996) *Complexity and creativity in organizations*. Oakland: Berrett-Koehler Publishers.

Stacey, R. (2001*) Complex Responsive Process in Organizations: Learning and Knowledge Creation*. London: Routledge

Stacey, R.D. (2003) *Complexity and group processes: A radically social understanding of individuals*. London: Routledge.

Stake, R. E. (1995) *The art of case study research* . Thousand Oaks, CA: SAGE Publications.

Steen, T., Nabatchi, T. and Brand, D., 2016. Introduction: Special issue on the coproduction of public services.

Stiefel, M.C. Riley, C.L. Roy, B. Ramaswamy, R. and Stout, S. (2016). *100 Million Healthier Lives measurement system: progress to date. 100 Million Healthier Lives Metrics Development Team Report.* Cambridge: Institute for Healthcare Improvement.

Sterman, J.D. (1994) Learning in and about complex systems. *System Dynamics Review*, 10(2-3), pp. 291-330.

Stroup. W. F. (1997) Webs of Chaos: Implications for Research Design, in R. A. Eve, S. Horsfall and M. E. Lee (eds.) *Chaos, Complexity and Sociology: Myths, Models and Theories*. Thousand Oaks: Sage Publications.

Sullivan, H. and Skelcher, C. (2002) *Working across boundaries: collaboration in public services*. Palgrave Macmillan.

Swieringa, J. and Wierdsma, A. (1992). *Becoming a Learning Organization: Beyond the Learning Curve.* Wokingham: Addison-Wesley

Taylor, F.W. (1911) *The principles of scientific management*. New York: Harper.

Taylor, M.J., McNicholas, C., Nicolay, C., Darzi, A., Bell, D. and Reed, J.E. (2013) Systematic review of the application of the plan–do–study–act method to improve quality in healthcare. *BMJ quality & safety*, pp. bmjqs-2013.

Teisman, G.R. (2008) Complexity and management of improvement programmes: an evolutionary approach. *Public Management Review*, 10(3), pp. 341-359.

Teisman, G.R. and Klijn, E.H. (2008) Complexity theory and public management: An introduction. *Public Management Review*, 10(3), pp. 287-297.

Teisman, G., van Buuren, A. and Gerrits, L.M. eds. (2009) *Managing complex governance systems*. London: Routledge.

Teisman, G. and Gerrits, L. (2014) The emergence of complexity in the art and science of governance. *Complexity, Governance & Networks*, 1(1), pp. 17-28.

Tjora, A.H. (2006) Writing small discoveries: an exploration of fresh observers' observations. *Qualitative Research*, 6(4), pp. 429-451.

Tomolo, A.M., Lawrence, R.H. and Aron, D.C. (2009) A case study of translating ACGME practice-based learning and improvement requirements into reality: systems quality improvement projects as the key component to a comprehensive curriculum. *Quality and Safety in Health Care*, 18(3), pp. 217-224.

Tosey, P., Visser, M. and Saunders, M.N. (2012) The origins and conceptualizations of 'triple-loop'learning: A critical review. *Management Learning*, 43(3), pp. 291-307.

Trade Union Congress (2013) *Making Co-Production Work: Lessons from Local Government*. London: Trade Union Congress.

Tsoukas, H. (2005) *Complex knowledge: studies in organizational epistemology.* Oxford: Oxford University Press.

Tsoukas, H. (1989) The validity of idiographic research explanations. *Academy of management review*, 14(4), pp. 551-561.

Tunstill, J. and Blewett, J. (2013) Mapping the journey: outcome-focused practice and the role of interim outcomes in family support services. *Child & Family Social Work,* 20(2), pp. 234–243

Uhl-Bien, M., Marion, R. and McKelvey, B. (2007) Complexity leadership theory: Shifting leadership from the industrial age to the knowledge era. *The leadership quarterly*, *18*(4), pp. 298-318.

Usher, R. and I. Bryant (1989). *Adult Education as Theory, Practice and Research: The Captive Triangle*. London: Routledge

Vaillancourt, Y. (2009) Social Economy in the Co-Construction of Public Policy. *Annals of Public and Cooperative Economics*, 80(2), pp. 275-313.

Van Manen, M. (1990) *Researching lived experiences*. Albany: State University of New York Press.

van Meerkerk, I., Boonstra, B. and Edelenbos, J. (2013) Self-organization in urban regeneration: A two-case comparative research. *European Planning Studies*, 21(10), pp. 1630-1652.

van Thiel, S. and Leeuw, F.L. (2002) The performance paradox in the public sector. *Public Performance & Management Review*, 25(3), pp. 267-281.

Vandenbroeck IP, Goossens J, Clemens, M. Foresight Tackling Obesities: Future Choices—Building the Obesity System Map. Government Office for Science, UK Government's Foresight Programme

Vargo, S.L. and Lusch, R.F. (2004) Evolving to a new dominant logic for marketing. *Journal of marketing*, 68(1), pp. 1-17.

Virginia Performs (2016) Measuring Virginia. Available from: http://vaperforms.virginia.gov/ (Accessed: 11 June 2016)

von Bertalanffy, L. (1968) *General systems theory*. New York: Braziller.

Von Hippel, E. (1994) "Sticky information" and the locus of problem solving: implications for innovation. *Management science*, 40(4), pp. 429-439.

Voorberg, W.H., Bekkers, V.J. and Tummers, L.G. (2014) A systematic review of co-creation and co-production: Embarking on the social innovation journey. *Public Management Review*, 17(9), pp. 1333-1357.

Waldrop, M. Mitchell (1993) *Complexity: the Emerging Science at the Edge of Order and Chaos.* New York: Touchstone.

Ward, S. C. (2011) COMMENTARY: The machinations of managerialism: New public management and the diminishing power of professionals, *Journal of Cultural Economy,* 4(2), 205-215.

Watzlawick, P., Weakland, J. and Fish, R. (1974) *Change: Principles of problem resolution*

and problem formation. New York: Norton.

Weber, M. (1922) *Bureaucracy in Classics of Public Administration.* Boston: Schafritz & Hyde.

Weiss, C.H. (1995) Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. *New approaches to evaluating community initiatives: Concepts, methods, and contexts*, 1, pp. 65-92.

Wenger, E. (1998) *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press.

Wenger, E. (2010) Communities of practice and social learning systems: the career of a concept. In *Social learning systems and communities of practice* (pp. 179-198). Springer London.

White, M., Wells, J.S. and Butterworth, T. (2014) The impact of a large-scale quality improvement programme on work engagement: Preliminary results from a national cross-sectional-survey of the 'Productive Ward'. *International journal of nursing studies*, 51(12), pp. 1634-1643.

Wheatley, Margaret J. (1994) *Leadership and the new science: Learning about organization*

from an orderly universe. San Francisco: Berrett-Koehler Publishers.

Wheeler, J. (2000) *Understanding Variation, The Key to Managing Chaos. 2nd ed.* Knoxville: SPC Press.

Weiss, R.S. (1995) *Learning from strangers: The art and method of qualitative interview studies*. New York: Simon and Schuster.

Wiener, N. (1948) *Cybernetics*. Paris: Hermann.

Wilson, W. (1887) The study of administration. *Political science quarterly*, 2(2), pp. 197-222.

Wimbush, E. (2011) Implementing an outcomes approach to public management and accountability in the UK—are we learning the lessons? *Public money & management*, 31(3), pp. 211-218.

World Health Organisation (2008) *Closing the Gap in a Generation: Health Equity through Action on the Social Determinants of Health*. Geneva: The World Health Organization.

Yazan, B. (2015) Three approaches to case study methods in education: Yin, Merriam, and Stake. *The Qualitative Report*, 20(2), pp. 134-152.

Yin, R. K. (2009) *Case study research: design and methods*. Sage publications.

Zeckhauser, R. and Shepard, D. (1976) Where now for saving lives? *Law and contemporary problems*, 40(4), pp. 5-45.

Zuckerman, D. (2013) Hospitals Building Healthier Communities. *The Democracy Collaborative at the University of Maryland, Report, Retrieved on*, 7, p.2014.

Appendices

Appendix A. Ethics approval form

sd/MM

7th April 2014

Max French

Dear Max

Understanding how Service Users and their Communities contribute to Public Service Outcomes

Thanks for your application to the School of Applied Social Science Ethics Committee for ethical scrutiny of your project. I can confirm that you have addressed the points raised and Chair's action has been taken to approve this proposal.

I wish you well with the study.

Yours sincerely

Margaret S. Malloch

Dr Margaret Malloch
Chair of Ethics Committee

Appendix B. Sample topic guide for first PM interview

The integration of the EYC within the CPP

o How the EYC has been adapted into CPP infrastructure

o Level of / satisfaction with senior support

o The role of the PM within the EYC

Response to Scottish Government infrastructure

o The use of Stretch Aims within strategic management / TOC development

o Perceptions of the Learning Sessions

o Perceptions of the Extranet

o Perceptions of the Workstreams

Improvement achievements

o Perceptions of the effectiveness of TOCs

- Numbers having scaled, failed, struggling, making progress

- Knowledge of TOCs achieving scale and perceptions of barriers

- Criteria for success in TOCs

o Perceptions of the effectiveness of the MFI

o Knowledge of sharing / learning from other CPPs and perceptions of barriers of this

Function of the EYC within the CPP

o Perception of value of the EYC

o Perception of the influence of Scottish Government

o Perception of the influence of ICS partnership / CPP

Strategy for moving forward

o Principal lessons learned

o Priorities for the EYC moving forward

Appendix C. Sample topic guide for final programme manager interview

Role of the PM

- o Change in the role compared with first interview

  - Technical / operational role

  - Networking / boundary spanning role

  - Strategic management within the CPP

Response to Scottish Government infrastructure

- o Reasons for engaging with Bootcamps / Pioneers / Improvement Advisors

- o Response to the Stretch Aims

- o Response to Key Change Areas

- o Perceptions of the Learning Sessions and Key Change Events

- o Perceptions of the Extranet

Improvement achievements

- o Changes in TOC development strategy

- Perceptions of the effects of Workstream failure

- Knowledge of TOCs achieving scale and perceptions of barriers

- Knowledge of sharing / learning from other CPPs and perceptions of barriers

- Focussed rumination on key achievements specific to each CPP

Function of the EYC within the CPP

- Perception of value of the EYC

- Perception of the influence of Scottish Government / CPP priorities

- Focussed discussion around key events in each CPP: e.g. collapse of workstreams, shifts of focus, ICS decisions

Strategy moving forward

- Perceptions of the main achievements of the EYC

- Perceptions of how the ambitions for the EYC within the CPP have shifted since initiation

- The future for the EYC and specific TOCs within the CPP

Appendix D. Sample topic guide for TOC leader interview

Context of project

- How would you describe your professional role? Would describe yourself as a manager, or a practitioner?
- What encouraged you to become involved with the EYC?

Improvement Project Details

- [Interviewer confirms what is known about project through document review]
- Can you describe a little bit about the history of the improvement project?
- Where are the main priorities of the project?
    - Have these changed?
- What's next for the project?

Idea generation

- What made you choose the aim?

[Interviewer probes participant to consider where ideas emerged from]

- What made you choose the [main changes participant has initiated to achieve improvement]?

[Interviewer probes participant to consider where ideas for changes emerged from]

Testing process

- Can you describe your engagement with the model for improvement?

- Do you feel the model for improvement encouraged or inhibited improvement?

[Interviewer probes: in what way?]

- How is the model for improvement viewed by your [colleagues / managers / practitioners]?

Scaling process

- Have you given any thought to how you might scale up your project?

[Do you find / Have you found] the model for improvement supports the scaling of improvements?

- Have you shared, or adapted anything from other Tests of Change?
- Have you spread learning from the project in other ways?

Stakeholder involvement

- How were practitioners / managers involved with the project?
  - How suitable were these roles?
- How did the project integrate with the wider service you provide?
  - Were you afforded the time and resources you needed to carry out the improvement project?

Engagement with EYC infrastructure

- Have the Stretch Aims or Key Change Areas been influential the project?
  - [IF SO – how have they been influential? IF NOT – why haven't they influenced you?]

- Are you interested in engaging with the wider Early Years Collaborative membership?

- What involvement has the Programme Manager had in the project?

- What engagement have you had with training, or Improvement Bootcamps?

- What has been your experience with the Learning Sessions / Extranet?

Appendix E: Observation protocol used for Workstream meetings

Workstream Observation Protocol

Workstream being observed:

Date:

Start time / end time:

Individuals present (total number and professional role):

Agenda item / topic of discussion	Researcher reflections and points of interest	Instance of collaboration observed

Appendix F. Summary of CPP 2 Home Team survey

Summary of responses

Total No. responses	Senior Manager / Leader	Operational Manager	Senior Practitioner	Practitioner	Other
22	45.45%	40.91%	9.09%	4.55%	0.00%

EYC Achievements

QUESTION	Strongly Disagree	Disagree	Neither agree nor disagree	Agree	Strongly Agree
EYC enables me to deliver quality improvement and change ideas in my organisation	0%	0%	27%	68%	5%
Vision of the EYC is widely and easily recognised	0%	64%	18%	59%	9%
Strong resistance in my team to embedding and applying quality improvement activities and change ideas	14%	45%	32%	9%	0%
Quality improvement activities are being facilitated in collaboration (people out with my team)	0%	14%	27%	50%	9%
Senior leaders & managers are supportive to enabling practitioners lead quality improvement activities	0%	0%	9%	73%	18%
EYC effectively communicates ongoing activities and developments	0%	5%	5%	78%	14%

Autonomy to develop quality improvement in my workplace	0%	0%	14%	55%	32%
Learning culture created in Fife to share and develop quality improvement	0%	0%	19%	72%	9%
Clear evidence is available that Fife EYC is contributing to improving outcomes for children and families	0%	5%	17%	64%	14%

Service user involvement

QUESTION	No	Not applicable	Yes – indirectly	Yes – directly
Children and families have influenced the development of quality improvement?	10%	15%	65%	10%
Your idea/s for improvement has drawn from interactions with children and/or their families?	5%	15.0%	25%	55%

Appendix G. Summary statistics for TOC population of CPPs 1-3

The following table and figures are composed based on TOC profiles across case study CPPs 1-3 in August 2015.

Table G1. TOC development by self-reported rating scale

	Total TOCs listed	Classed as active testing (rating 2 or higher)	Classed as having achieved improvement (rating 3 or higher)	Classed as having achieved 'significant' improvement (rating 4 or higher)
CPP 1	16	9	7	0
CPP 2*	34	34	31	15
CPP 3	18	12	6	2
Total CPP sample, August 2015	68	57 (84%)	43 (63%)	16 (24%)

*Note that CPP 2's figures are driven partially by a different interpretation of the rating scale

Figure G1. Categorisation of case study CPP TOCs by service context

Pie chart legend:
- Other areas
- Nurseries
- Primary Schools
- Allied Health Services
- Community Services
- Health and Development Promotion

Pie chart values: 24%, 16%, 12%, 15%, 18%, 15%
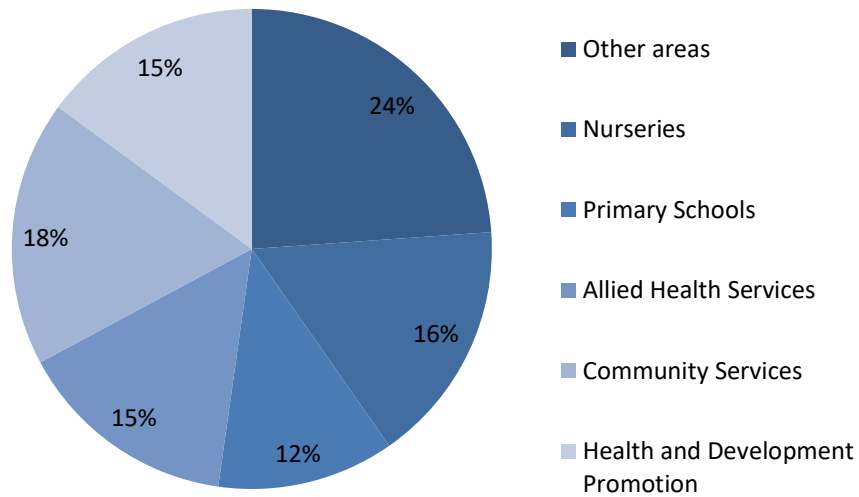
Figure G2. case study CPP TOC alignment to KCAs

Bar chart categories and values:
- Parenting skills and family engagement: 18
- Attachment, child development and learning: 16
- Continuity of care in service transitions: 7
- Health and Wellbeing: 7
- Early support for pregnancy and beyond: 6
- Addressing child poverty: 5
- Play: 4
- 27-30 month child health review: 2
- Not given: 2
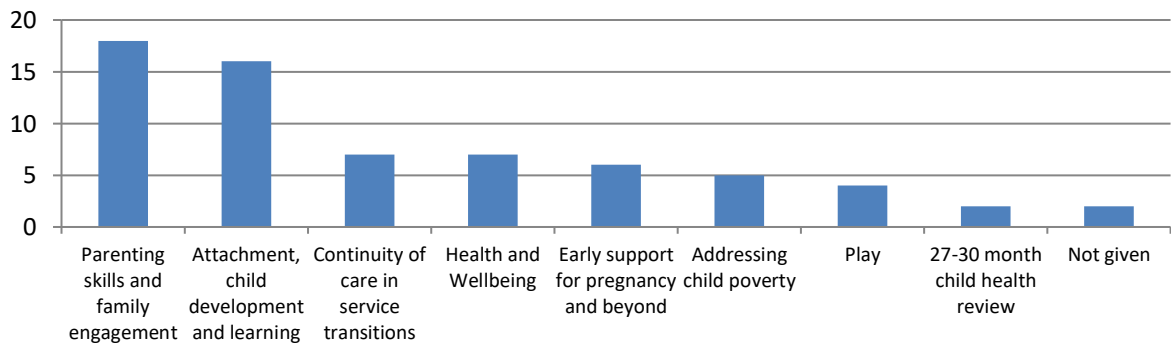
Appendix H. Stretch Aim Driver Diagrams

WORKSTREAM 1 (conception to 1 year)

Theory of what drives infant mortality

Theory of what actions will reduce infant mortality

Aim	1º	2º

Reduce infant mortality

Societal Issues
- Poverty
- Quality Of Home Environment
- Domestic Abuse & Violence
- Workforce Issues
- Transport, Community Capacity & Cultures
- Access To Services
- Employment

Post-birth actions
- Attachment
- Health
- Parenting skills

Pre-birth actions
- Smoking / Alcohol & Drug Misuse
- Nutrition
- Mental health & wellbeing
- Pre-Conception

Detailed aim:

To ensure that women experience positive pregnancies which result in the birth of more healthy babies as evidenced by a reduction of 15% in the rates of:

- stillbirths (from 4.9 per 1000 births in 2010 to 4.3 per 1000 births in 2015)

- infant mortality (from 3.7 per 1000 live births in 2010 to 3.1 per 1000 live births in 2015)

Actions:
- Improved teamwork, communication, skills and collaboration
- Improved money management
- Improved rate of breastfed babies
- Quicker diagnoses of Neonatal Abstinence Syndrome
- Improved leadership, culture and planning
- Improved family centred response
- Improved stability / permanence for LAC
- Improved identification
- Improved joint working
- Improved management and quality of care
- Improved sharing of information
- Improved access
- Identification & reasons for current resilience

Version: 10/04/2013

385

WORKSTREAM 2 (1 year to 30 months)

Theory of what drives developmental milestones

Theory of what actions will ensure developmental milestones are reached

Aim	1°	2°

Aim

Children have all the developmental skills and abilities expected of a 27-30 month old

Detailed Aim:

85% of all children within each CPP have reached all of the expected developmental milestones at the time of the child's 27-30 month child health review by end-2016

1°

Societal Issues

Child's physical & mental health and emotional development

Carer's physical & mental health and skills

2°

Poverty

Quality Of Home Environment

Domestic Abuse & Violence

Workforce Issues

Transport, Community Capacity & Cultures

Access To Services

Employment

Health

Attachment

Early Learning & Play

Additional Support

Level of education

Misuse of alcohol & drugs

Nutrition

Disabilities & Mental health

Parenting skills & knowledge

Improved teamwork, communication, skills and collaboration

Improved money management

Improved child's dental health

Improving child nutrition

Improving brain development and physical play

Improved family centred response

Improved stability / permanence for LAC

Improved early identification

Improved joint working

Improved sharing of information

Improved management, planning and quality of services

Improved leadership, culture and planning

Identification & reasons for current resilience

Version: 06/03/2013

386

WORKSTREAM 3 (30 months to start of primary school)

Theory of what drives developmental milestones

Theory of what actions will ensure developmental milestones are reached at the start of primary school

Aim   1º   2º

Children have all the developmental skills and abilities expected at the start of primary school

Detailed Aim:

90% of all children within each CPP have reached all of the expected developmental milestones at the time the child starts primary school, by end-2017

Societal Issues

- Poverty
- Quality Of Home Environment
- Domestic Abuse & Violence
- Workforce Issues
- Transport, Community Capacity & Cultures
- Access To Services
- Employment

Child's physical & mental health and emotional development

- Early Learning & Play
- Health
- Attachment
- Additional Support

Carer's physical & mental health and skills

- Level of education
- Misuse of alcohol & drugs
- Nutrition
- Disabilities & Mental health
- Parenting skills & knowledge

- Improved teamwork, communication and collaboration
- Improved uptake of benefits
- Improved child's dental health
- Improving child nutrition
- Improving brain development and physical play
- Improved family centred response
- Improved stability / permanence for LAC
- Improved identification
- Improved joint working
- Improved management, planning and quality of services
- Improved sharing of information
- Improved leadership, culture & planning`
- Identification & reasons for current resilience

Version: 06/03/2013

387

WORKSTREAM 4 – DRIVER DIAGRAM

Aim
90% of all children in each CPP area will have reached all of the expected developmental milestones and learning outcomes by the end of Primary 4 by end-2021

Primary Driver
Societal Issues

Secondary Driver(s)
Poverty   Home environment   Domestic Abuse
Impact of inequalities   Media Influences
Access to wider opportunities
Employment /Workforce issues
Accessibility and affordability of childcare
CYP implementation

Primary Driver
Child's physical, cognitive, mental health & emotional development (Individual child's characteristics/profile)

Secondary Driver(s)
Safe & secure home   Access to quality Early learning   Play
Child's & family aspirations   Health   Attachment
Ability to form relationships   Participation in wider opportunities
Workforce knowledge, skills & understanding   Cultural activities
Disabilities
Child resilience

Primary Driver
Carers'/Parents' physical & mental health & skills

Secondary Driver(s)
Safe & secure home   Family resilience   Adequate income
Family aspirations   Level of education   Substance misuse
Disabilities & mental health   Parenting skills & knowledge
Nutrition/physical disabilities   Family relationships

Primary Driver
Learning environment
Achievement of Es and Os at relevant CfE level
Standards of delivery for meeting learners needs

Secondary Driver(s)
Home learning environment   Parent's engagement
Parental involvement with services   School learning environment
out of school childcare, learning and play   Support with homework

Primary Driver
Community environment

Secondary Driver(s)
Transport   Access to wider opportunities   Access to services
Children a key part of the community
Parents on school councils/ committees
Children's views sought and acted on in developing local resources
Space to play   Use of schools and grounds community assets

WORKSTREAM 4 (Leadership)

Theory of what drives leadership support

Theory of what actions will ensure leadership support

Aim   **1º**   **2º**

| Build commitment with partners to focus on delivery |
| CPPs communicate the EYC with enthusiasm and consistency |
| Leaders illustrate how users are included in design, improvement, and delivery of Early Years |

Early Years Collaborative is a strategic priority & underpins all policy planning and operational activity

Provide the Leadership System to support quality improvement across the Early Years Collaborative

| Leaders facilitate change by cultivating innovation from intelligence, insights and wisdom of people working together |
| Leaders demonstrate their ability to set direction and engage and mobilise staff to constantly improve quality of service |
| Leaders can describe how they personally maintain early years focus within their working environment |

Early Years Collaborative values, culture and behaviours are modelled by all leaders at all levels

| Early years executive and operational leads are identified |
| Measurement plan and priorities are established and triangulation with other key data |
| Spread plan is in place for core and innovative work |
| Strategy for capturing, celebrating and spreading innovation |

Infrastructure to support delivery of Early Years Collaborative

| Establish an EYC Implementation Committee |
| Ensure a feedback mechanism for issues raised in Walk-rounds |
| Ensure the development of a measurement system used to understand and drive quality indicators |
| Assign a senior leader to each improvement area (Workstreams 1-3 and measurement) |
| Establish Programme Management and remove barriers |
| Meet regularly with the Implementation Committee to track progress and remove barriers |
| Display data that depicts progress towards aim |
| Ensure that the senior team participates in Walk-rounds |
| Place quality issues at the top of senior leader meeting agendas |
| Add Early Years Collaborative and outcomes to the CPP agenda |

Detailed Aim:

Timely delivery of all three workstream "stretch aims"

Version: 06/03/2013

389

Appendix I. EYC key event timeline

Leadership workstream established with associated driver diagram

Improvement Advisors are recruited to 9 month training programme

Development of 'Stretch Aims' 1-3 outcome measures and associated 'Driver Diagrams'

Most CPPs have set up dedicated staff charged with carrying forward EYC activity

Half of all CPPs initiate pioneer sites; 40 are enlisted by LS5

Practice Development Team disbanded and 'Improvement Advisors' assume role

Policy Team engages Local Authority Chief Executives and all 32 CPPs sign up to delivery of the EYC

CPPs set up 'Workstreams' and begin to implement Tests of Change

Key Change Areas are developed and finalised. Recruitment for 'Pioneer sites' begins

Raising Attainment for All (RAFA) collaborative Launched

650 Tests of Change reported to be in operation

Scottish Government move toward unbadging the EYC institute the Children and Young People's Improvement Collaborative

2013

2014

2015

January Learning Session 1

October Learning Session 3

June Learning Session 5

March Learning Session 7

1st October 2012 EYC Launch Event

May Learning Session 2

January Learning Session 4

October Learning Session 6

November Joint EYC / RAFA Learning Session 8

391