Quality Improvement Interpreted as a Complex Adaptive System: Implications and Opportunities

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at Massey University Palmerston North, New Zealand

William James Wilson

2024

Abstract

The effectiveness of quality improvement (QI) methods in healthcare has been challenged, especially under circumstances of high complexity. This thesis examines the implications for quality improvement if complex socio-technical systems such as healthcare are interpreted as complex adaptive systems (CAS).

The research followed a mixed-method design. Informed by the complex systems and quality management literature, a conceptual model for quality improvement within CAS was developed — the *complex quality improvement network* (CQIN). An agent-based simulation model was then used to establish the plausibility and face validity of the model constructs and their interaction. Thematic analysis and crisp-set qualitative comparative analysis (QCA) were then used to examine the evidence for CQIN constructs within published quality improvement case studies. One applied case study was also conducted for deeper insight into the practical difficulties of interpreting a real-world quality improvement project as a CAS. Finally, the findings of the simulation modelling and the secondary data analysis were integrated into a Bayesian network model.

Empirical evidence, in the form of consistency across cases and coverage within cases, was found for eleven of the twelve CQIN constructs. Multiple sets of sufficient conditions for reported improvement success were identified across cases. These sets were minimised to four strategies for successful quality improvement; i) strengthening agent network communication paths; ii) building shared understanding of problem and context amongst networked agents; iii) increasing problem-solving effectiveness; and iv) improved system signal integration. If the evolutionary foundations for CAS are in some way inhibited, the likelihood of quality improvement success is reduced.

Healthcare quality improvement can be plausibly simulated using fundamental CAS principles. The first contribution to quality improvement discourse is the CQIN model, a CAS model of change applied specifically to quality improvement. A second contribution of this research is a complex quality improvement risk assessment model, the CQIN Bayesian Network. Practitioners can use this model to examine and test identified CAS-informed improvement strategies. The individual CQIN constructs make a third contribution by providing new categories of causal factors for the comparison of disparate quality improvement case studies.

Acknowledgments

This thesis would not have been possible without the support and inspiration of many people. I will start with my supervisors, Dr Nihal Jayamaha, Dr Kudakwashe Dube, Dr Scott McLachlan, and Dr Nicola Brown. All have provided regular guidance and mentoring, cheerfully accepting the challenge of teaching me to think and express myself in a scientific manner. Nihal, thank you for your unstinting support and encouragement over many years. This research would not exist without you. Kuda, thank you for your disciplined logic and your astute feedback despite the geographical challenges. Scott, thank you for taking on the supervision and mentoring of someone so green. Throughout the entire research process, you have provided me with ideas, references, contacts, and technical examples to explore. You also worked out that I can cope with blunt assessments when my efforts were below required standards. I hope I can shake your hand and say thank you in person one day. Nicola, thank you for being willing to join the team at short notice and providing calm guidance in the final months of work.

I must thank Massey University for the opportunity to undertake doctoral study and for the generous scholarship that allowed the idea to become a reality. I would have been unable to continue without this generous support. This thesis is the capstone on my involvement with the Massey Quality Systems, Logistics and Supply Chain team, and I cherish the experience. Thanks again to Nihal and also Dr Nigel Grigg for the inspiration and the elevation of quality as a topic worthy of serious study.

Next, I would like to acknowledge Dr David Kerr, and Dr Kathleen Potter and Bruce Penny for their enthusiasm and guidance as I got started. I finished up a long way from those early explorations, but somewhere very interesting, nevertheless. Kathleen, I valued our regular discussions immensely and your scientific rigour shows me how high that bar really is. Bruce, thank you for your friendship and wisdom and for valuing my work. You might even find some consideration of the art of leadership within these pages.

To my Te Whatu Ora Waitaha managers Richard Hamilton and Lynn Davies, thanks for all the encouragement, practical support, and freedom to follow my own path. I would also like to acknowledge Megan Hopper, Dr Richard Floyd, and Dr Saxon Connor for the many enjoyable conversations we have had about quality and how to reimagine it.

Finally, to my wife Helen, thank you for your patience, for believing in me, for putting up with me hidden away in my study, the long list of deferred house maintenance, and for enduring jargon-filled conversations about complex systems. That is a debt I cannot repay.

This thesis is dedicated to my uncles, Dr Logan McLennan and Dr Colin McLay. You have been my role model scientists, inspiring me with your curiosity, intellect, compassion, and humour. And of course, to Noel and Bruce, who gave me a such a good start in life that anything seemed possible.

Related publications

Wilson, W., McLachlan, S., Dube, K., Potter, K., & Jayamaha, N. (2023). Uncertainty, emergence and adaptation: A complex adaptive systems approach to quality improvement. *The Quality Management Journal*, *30*(3), 168–186. https://doi.org/10.1080/10686967.2023.2211287

Table of Contents

List of Acronyms

ABM	Agent-Based Model
ARC	Aged Residential Care
BN	Bayesian Network
CAS	Complex Adaptive System
CFIR	Consolidated Framework for Implementation Research
CQIN	Complex Quality Improvement Network
DAG	Directed Acyclic Graph
EBM	Evidence-Based Medicine
GST	General Systems Theory
LHS	Learning Health System
MAC	Medications Advisory Committee
NPT	Node Probability Table
NZ	New Zealand
PDSA	Plan-Do-Study-Act
QCA	Qualitative Comparative Analysis
QI	Quality Improvement
QIC	Quality Improvement Collaborative
SIR	Susceptible-Infected-Resistant
SPC	Statistical Process Control
SSM	Soft Systems Methodology
UML	Unified Modelling Language

List of Tables

List of Figures

Part 1: Background

Chapter 1 Introduction

1.1. Research Motivation

Quality performance indicators for healthcare systems internationally appear to have plateaued at sub-optimal levels. According to one recent assessment, only around 60% of care is based on the latest evidence, up to 30% of healthcare spending is wasted, and approximately 10% of patients suffer adverse events as a result (Braithwaite, 2018). For Braithwaite, this level of performance is embedded and has resisted decades of improvement effort and system restructuring. In the New Zealand (NZ) health context, equality of access, service, and outcome across demographic groups, timely access to care, and workforce retention are especially problematic (Goodyear-Smith & Ashton, 2019). At the time of writing, New Zealand is once again attempting to address these issues and is undertaking ambitious whole-of-system healthcare reforms (Pae Ora Healthy Futures Act, 2022). New Zealand's healthcare delivery problems are seemingly difficult to solve but as a research community we care enough to keep trying to improve the situation (Horn & Gorman, 2021). We are all consumers of health systems and in most cases, indirect funders via taxation. Effective and sustainable health systems are a *sine qua non* for flourishing societies and worthy of ongoing improvement effort. The motivation for this research is the desire to contribute to improved healthcare delivery, coupled with a clear line of enquiry to investigate — the potential of complexity-informed QI.

1.2. Research Scope

At a high level of generalisation, the research topic can be described as 'problem solving under conditions of complexity'. What makes something complex? The Concise Oxford English dictionary defines complex as *consisting of many parts* and *not easy to understand; complicated* (Stevenson & Waite, 2011, p. 293). This simple dictionary definition suffices to highlight two important complexity concepts: (i) a high number of constituent parts; and (ii) uncertainty about what exactly is going on.

The provision of healthcare services is undoubtedly complex (Covvey, 2018). The assertion that healthcare is complex is easy to defend. Human behaviour is complex (Schill et al., 2019). Human physiology is complex (Bashan et al., 2012). Therefore, the healthcare systems we create must also be complex. Faced with very difficult problems, it is natural for improvement practitioners to try and expand their toolkit. A sensible way

to increase solution options is to draw on the experience of other knowledge domains, such as complexity science. It is no surprise that healthcare researchers have actively explored complexity ideas and references to complexity are now commonplace in the QI literature (Braithwaite et al., 2018; Dixon-Woods & Martin, 2016; Khan et al., 2018).

The research scope targets QI as the area of attention within healthcare more generally. From within QI, the precise focus is the convergence of quality improvement and complexity; specifically, the need for a fully complex adaptive systems (CAS) model of quality improvement. Such a model will help answer two fundamental questions. First, is it valid to consider intentional, goal-oriented improvement within socio-technical systems as a form of CAS? Second, if the characterisation is valid, what do improvement methods built on CAS principles look like?

1.3. Background: Complexity and Complex Adaptive Systems

The labels *complexity*, *complexity science* and *complexity theory* are umbrella terms for overlapping subject areas with historical origins in physics, mathematics, biology, ecology, psychology, sociology, engineering, and computer science. These origins can be grouped into five fundamental intellectual traditions within the complexity sciences: (i) *Dynamical systems theory, (ii) Systems theory, (iii) Complex systems theory, (iv) Cybernetics and (v) Artificial intelligence* (Castellani & Gerrits, 2021).

The varied origins of these complexity traditions have led to multiple interpretations of complexity (Adami, 2002). When complexity principles are applied into novel areas of study, the tools and approaches of the different source traditions naturally tend to persist. Practitioners need to undertake careful research before determining which approach or tool is best suited to the new context of interest. Whilst some requirement to learn the territory is to be expected, the combination of inherited assumptions and the need to master new analytical tools can be daunting for practitioners. Exactly which complexity concepts are relevant? Do they apply to the specific context under consideration? As a result, there are continued calls in the literature for integrated, theory-driven approaches to bind these diverse efforts together into generalisable and scalable solutions applicable to healthcare QI (Braithwaite et al., 2018; Davidoff et al., 2015; Greenhalgh & Papoutsi, 2018).

The idea of *complex adaptive systems* (CAS) emerged in the 1960s from within the *complex systems theory* research tradition. CAS ideas have subsequently become a

dominant area of complexity thinking across many research domains. Google Scholar returned 106,000 results on 29th January 2024 for the search query *complex adaptive systems*. Why have CAS ideas resonated so strongly with researchers in different knowledge areas? One potential reason is that complexity is so pervasive across many problem-solving domains it is hard to avoid. Another driver may be the fact that early CAS scholars made strong claims for generalisability beyond the origins of CAS in the physical and life sciences (Holland, 2006; Kauffman, 2000; Miller & Page, 2007; Mitchell, 2009).

The suitability of adopting a CAS framework to explain sophisticated socio-technical improvement endeavours beyond natural science settings has largely gone unchallenged (several critiques are noted in Chapter Three). *Socio-technical* is used here in the sense of the dynamic interactions between social behaviour and the technical components of a 'system', including information technology systems and organisational constraints such as policies and business rules (Appelbaum, 1997; Trist, 1981). Given the prevalence of seemingly complex situations, a major appeal of CAS might also be simply one of recognition; CAS descriptions just seem to align well with the problems researchers and practitioners find difficult to solve. Examined critically, this 'recognition' of CAS phenomena may be found to be confirmation bias, an attractive metaphor rather than a valid explanation.

Whether it is a useful metaphor or something more fundamental, the sense of recognition with CAS concepts amongst healthcare professionals is understandable. Healthcare often involves (i) very high numbers of 'agents'; (ii) high diversity amongst the participating agents/agent groups, including patients, nurses, doctors, allied health clinicians, family members, managers, administrators, analysts, software and technology providers, and politicians at multiple levels of government; (iii) unpredictable causal paths; (iv) poor information flow; and (v) lack of system stability or equilibrium. Healthcare change processes are seemingly at odds with highly prescriptive interventions. Outcomes emerge from multiple interactions, there is a resistance to centralised control, and self-organisation is preferred. The characteristics above are exactly what CAS proponents describe as CAS features (Buckley, 2017; Holland, 2006; Mitchell 2009; Turner & Baker 2019). It is a rational response for change agents, when confronted with difficult problems, to question their methods and look to an approach that not only better describes their perceived context but also suggests alternative means to bring about

change. If healthcare systems are not merely difficult bureaucracies, but are indeed *complex adaptive systems*, then it is essential to thoroughly investigate the potential insights complexity principles may offer our improvement efforts.

1.4. Research Aim

The healthcare QI literature suggests that conventional QI methods may have limited success improving health quality outcomes in the presence of high levels of complexity. This research investigates what may be missing from existing QI methods when working within high levels of complexity. Complexity-informed QI models will enable easier identification and handling of barriers to QI work than is possible with conventional QI methods. This will increase the level of success in quality outcomes in complex organisational systems such as national healthcare systems.

1.5. Research Question

Research Question

What system factors influence the success or failure of healthcare QI activity within complex adaptive systems?

To answer the research question, this research develops and tests theory in the form of a CAS model of change within healthcare QI.

1.6. Research Hypotheses

Four interlinked hypotheses will be examined to test a CAS model of socio-technical change and answer the research question. An initial premise is stated first as it underpins the entire healthcare as CAS discourse. While the evidence for regarding complex socio-technical organisations as CAS is carefully considered, this study cannot prove or disprove this premise.

Starting Premise: Complex human socio-technical organisations function as CAS.

H1. Change within socio-technical CAS will be consistent with CAS principles.

H2. Conditions for successful change within CAS can be identified.

H3. Within socio-technical CAS, system-level change in the form of goal-directed improvement is achieved through successfully coordinated updates and alignment of agent schema.

H4. Updating and aligning agent schema is achieved within the constraints of a limited set of agent behavioural rules.

The study commences with the logical implication following the starting premise. If the premise is true, then CAS explanations of intentional change within complex human socio-technical systems should exist. The research will develop, validate, and verify a specifically CAS explanation of change within real-world examples of complex problem solving. If the starting assumption is accepted, the first hypothesis describing intentional change follows — *change within socio-technical CAS will be consistent with CAS principles* (H1). The literature also suggests that *conditions for successful change within CAS can be identified* (H2) and any support for this claim would be a practical contribution of the research.

H1 and H2 are broad statements, containing multiple CAS principles which can be tested in many ways. H3 proposes that role of the agent internal schema is essential in generating emergent, system-level phenomena. H4 bounds any alignment of agent schema to a limited set of agent behavioural rules. H3 and H4 are core tenets of CAS behaviour, described in more depth in Chapters 2 and 3.

1.7. Research Objectives

The research has five objectives:

1. To investigate the role of system factors as possible components of change within CAS.
2. To examine the interaction of system factors within CAS.
3. To explore the relative influence of system factors as components of change within CAS.
4. To investigate factors that may constrain change within CAS.
5. To validate and verify the results using simulation and secondary data from published case studies in healthcare QI.

Collectively, research objectives 1-5 support the testing of Hypotheses 1-4, and each investigation of the hypotheses contributes to answer the research question. Objectives 1-4 have a difference in emphasis: the function of system factors; the interaction of system factors; and, the relative influence including constraining effects. Research Objective 5 ensures a transparent and valid process is followed, and that empirical evidence supports the hypotheses.

1.8. Outline of Research Method

This research investigates quality improvement within complex adaptive systems using simulation modelling and secondary data analysis to overcome the difficulty of directly measuring complex systems. Agent-based modelling (ABM) is used for the initial experimental analysis, due to its ability to easily represent CAS phenomena, in this case large networks of agents, defined agent rules, and a simple diffusion (sharing) process amongst the agent network (Borrill & Tesfatsion, 2011; Laubenbacher et al., 2007).

To see how well the simulation model represents real-world activity, verification of the findings will be examined by a secondary data analysis referenced back to the case study portion of the literature, triangulating theory and practice. The secondary data analysis is presented using qualitative comparative analysis (Ragin, 1999). The conceptual model provides a unifying language across 90 published QI cases studies where the study authors have either described or consciously reflected on aspects of complexity within their work.

A second computational model will then be developed using Bayesian Networks (BN). QI is inherently Bayesian in practice, in that practitioners continually learn from experience and update their hypotheses. Updating quickly and repeatedly as new evidence is revealed is key for working within complexity. Structurally, BNs allow a large set of interconnected variables to be managed in a transparent set of dependencies. BN also have an ability to cope with sparse, noisy or ambiguous data, and allow for evidence from both empirical data and expert opinion. Agents in complex systems are subject to incomplete levels of knowledge driven by the number of system elements and the interactions between these elements. The result is an environment where the consequences of deliberate actions can be unpredictable. Cause and effect may not be clear. BN can support agents make this epistemic uncertainty transparent via probabilistic reasoning (Fenton & Neil, 2019; Pearl & Mackenzie, 2018). Unpredictability arising from random chance can similarly be modelled using probability.

In summary, BN are an ideal analysis tool for complex problem solving (Fenton & Neil, 2019). Reproducing the underlying conceptual model via two unrelated modelling methods enhances concept validity by ensuring that results are not merely artefacts produced by the modelling environment or construction. The two operational models are built on different modelling ‘substrates’ and so are not exact replicas. However, alignment

of results and the flexibility demonstrated contributes to confidence in the underlying conceptual model's usefulness.

The research uses simulation experiments and secondary data to identify a plausible adaptive process within socio-technical CAS. The goal of the research question and hypotheses is not empirical validation of existing or novel QI theory. Rather, this is a contextually bound (healthcare QI) search. An equivalent, non-biological, adaptative process is a prerequisite to taking seriously the assertions within CAS and QI literature that human socio-technical systems are literally CAS (Braithwaite, 2018; Holland, 2012; Kahn et al., 2018; Kauffman, 2000).

1.9. Expected Research Contribution

This research anticipates three theoretical and practical contributions to quality improvement scholarship and practice:

1. A conceptual model that provides a specifically CAS explanation of QI. Such an explanation is a pre-requisite for the valid adoption of CAS principles in novel social system contexts. In addition to supporting theory, a CAS explanation of change identifies new improvement strategies for practitioners based on the mechanisms of CAS.
2. A ready to use, practical tool in the form of a complex quality improvement risk assessment model. Improvement practitioners can use this model to examine and test identified CAS-informed improvement strategies.
3. Systematic literature reviews reveal a wide scattering of themes and interpretations within the general description of 'complexity' (Thompson et al., 2016; Turner & Baker, 2019). The conceptual model constructs provide unifying categories for complexity-referencing QI activity.

1.10. Potential Impact

Within a CAS, change in system-level outcomes emerges from agent behaviour. So-called 'simple' rules and behaviour drive the desired higher order effects (Holland, 2012; Khan et al, 2018). A CAS explanation of QI will do more than address a gap in the healthcare QI literature. It will change the way QI is undertaken. If the CAS explanation is valid, then new ways of working will be required to bring about intentional change within CAS. Much more attention will need to be given to the network of cooperating

agents — their connections, opportunities to share information, internalised world views, and perceived agency.

1.11. Thesis Outline

This thesis is presented in three parts as outlined in Figure 1.1.

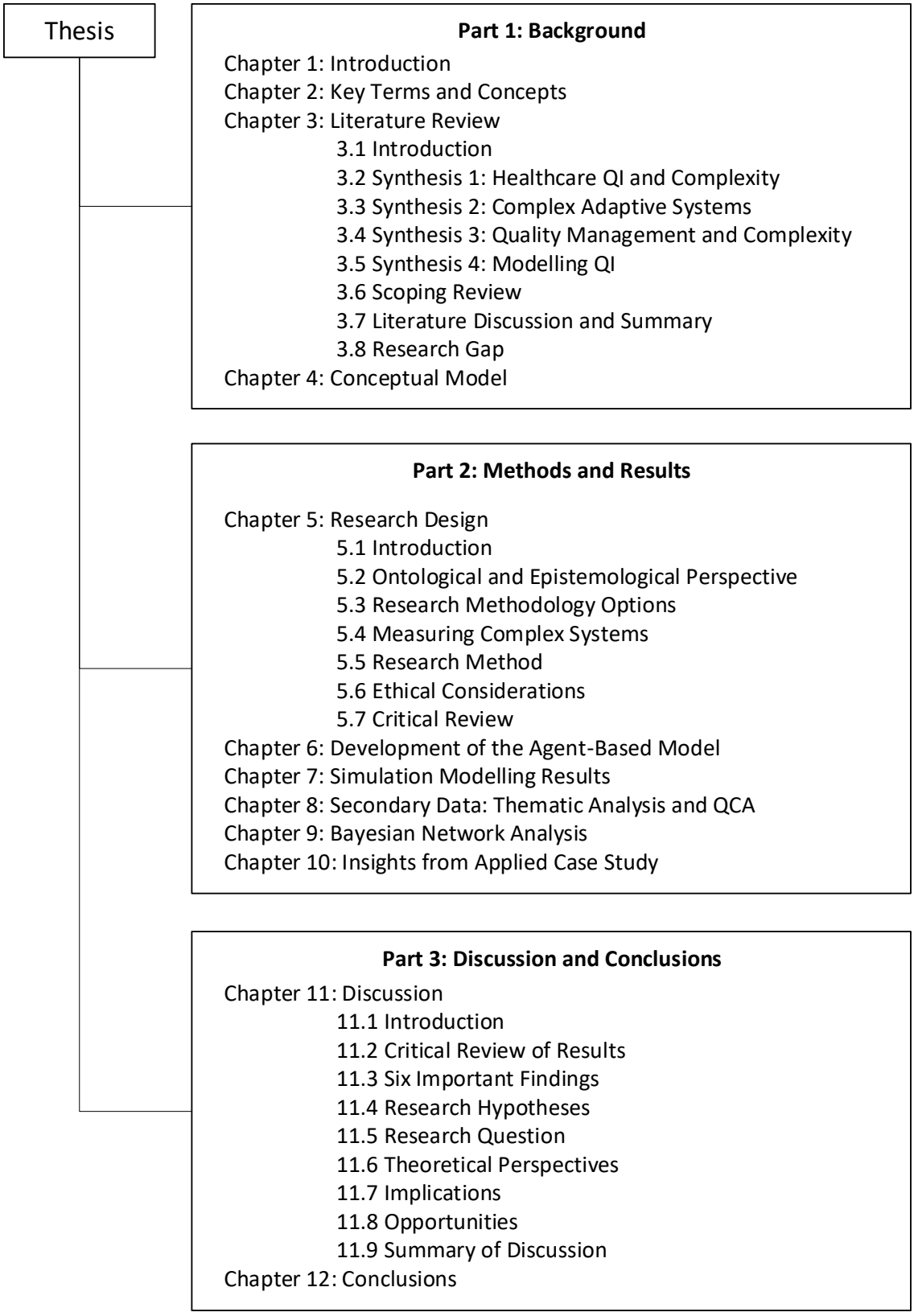


Figure 1.1: Outline of Thesis Chapters

The first part of the thesis contains four chapters introducing the research and the state of current discourse on the topic. The research context is established first via Chapter

1: Introduction. Key terms and concepts are introduced in Chapter 2 and the literature review follows in Chapter 3. The literature review is made up four separate syntheses of the literature, covering:

i. Healthcare and complexity thinking
ii. Complex adaptive systems.
iii. Quality management and complexity.
iv. Modelling approaches to quality improvement.

The literature review also contains a scoping review focused on the specific intersection of healthcare QI and complexity. In addition to important contributions to the syntheses above, the scoping review identifies a body of QI improvement project case studies suitable for empirical validation of the conceptual model. The literature review concludes with a discussion and summary of the major findings and a description of the research gap. Following on from the literature review and completing Part 1, Chapter 4 then introduces the conceptual model, including a consideration of the underlying theory operationalised by the model.

Part 2 of the thesis contains the research design, methods and results. This commences with the research design and method considerations in Chapter 5. The detailed research steps are set out, beginning with the research question and covering the use of simulation experiments, qualitative insights from an applied case study, and validation and data triangulation using secondary data. Chapter 6 presents the technical development of the agent-based operational model. Chapter 7 then examines results and insights from the agent-based simulation studies. The thematic analysis of the secondary data is set out in Chapter 8. As part of the research design, the published cases studies are used as secondary data to assess the conceptual model. The data are published QI studies where the researchers are consciously applying or referencing complexity principles. This published QI activity is used to validate and critique the conceptual model. In other words, does the conceptual model adequately describe and explain real-world quality improvement? Chapter 9 then brings the conceptual model, simulation model findings and the secondary data analysis together into a Bayesian Network Model. This model incorporates the separate research threads to build a predictive tool for complex quality improvement. Chapter 10 examines the practical considerations for operationalising the CQIN principles and models, using the lessons from the qualitative case study.

Part 3 of the thesis begins with Chapter 11 and the discussion of the research results in terms of both the theoretical implications and the practical opportunities. Chapter 12 concludes the thesis, summarising the research, the important themes and why they matter, the limitations of the research and the potential for future research.

Chapter 2 Key Terms and Concepts

2.1. Introduction

The language of complexity introduces a mixture of technical terms such as *non-linear* and *emergence,* together with domain-specific usage of everyday words such as *system* and *learning*. A contextual definition for each of the following terms is necessary for this study: *adaptation, agent, agency, agent network, coevolution, complex adaptive system, cooperation, emergence, environment, information, learning, measures of complexity, non-linear, problem solving, problem space, quality improvement, rejection, schema, shared schema, system signals, uncertainty*. Given that the research is investigating a CAS explanation of healthcare QI, these two concepts are defined first, then the remaining terms are explained and defined in the context of the research. The description of CAS also incorporates the important definitions of *adaptation* and *learning* as applied in this context.

2.2. Definition: Complex Adaptive Systems

Sociologist Walter Buckley is credited with introducing the term *complex adaptive system* (CAS) in 1968. Buckley defined CAS as:

> *"…a complex of elements or components directly or indirectly related in a causal network, such that at least some of the components are related to some others in a more or less stable way at any one time. The interactions may be mutual or unidirectional, linear or non-linear or intermittent, and varying in degrees of causal efficacy or priority."* (Buckley, 2017, p. 493)[1]

Buckley sought to enhance his field by looking closely at ideas from General Systems Theory (GST) and Cybernetics. From the beginning of CAS thinking, a serious effort was made to transfer ideas from technical to social knowledge domains. It was a 'pull' exploration, in the sense that Buckley and colleagues were seeking new ideas to advance their field, as opposed to having them imposed from outside. As von Bertalanffy conceived GST, and as Buckley clearly viewed it, GST aspired to be a genuinely universal theory, applicable to biological, non-human and human adaptive systems (von Bertalanffy, 1972; Buckley, 2017).

The characteristics of *multiple elements, complex interactions, instability, causal networks and relationship networks* are identified in Buckley's definition. This serves as

[1] First published in 1968

a simplified description of a 'complex system', but it does not describe how the system functions. Buckley, von Bertalanffy and Ashby all sought to describe how organisms and systems functioned and ultimately survived or persisted (Ashby, 1956; Buckley, 2017; von Bertalanffy, 1972). The need for a system to change in some way adds the notion of *adaptive* into the complex system. Mitchell (2009, p. 13) describes adaptation as *a change in behaviour to improve chances of success — through learning or evolutionary processes.* In the socio-technical context, learning is the applicable process. Learning is the acquisition of new information that is used to update the internal mental models of agents. Learning may be tactical (task-focussed) or strategic, namely, determining how to best resolve a problem (Anderson, 2005).

Adaptive systems are able learn from and adjust to their environment. At one end of a continuum, this may be achieved via a direct feedback mechanism that adjusts components within a range of defined parameters — homeostasis in physiology or a servo-controller within an engineered system. Further along the change continuum, vastly more sophisticated responses will be required. Here, Buckley and colleagues highlight the importance of detecting and responding to *system signals*:

> *"A learning net functions as a society, in this view, to the extent that its constituent physical parts are capable of regrouping themselves into new patterns of activity in response to changes in the net's surroundings, or in response to the internally accumulating results of their own or the net's past".* (Karl Deutsch, quoted in Buckley, 2017).

System adaptation via learning adds *Learning networks*, *feedback*, and *system history* to the list of characteristics/features of CAS. Evolutionary principles underpin the thinking, in the sense that the system must learn from and adapt to environmental changes in order to survive. This in turn requires some form of mapping the environment into the system (Buckley, 2017, p. 492). Thus, genetic coding in biological systems and mental models within socio-technical systems. Also necessary for adaptation is the ability to generate options for change. Here, Ashby's (1956) *law of requisite variety* forms part of Buckley's argument. Simply put, any regulatory mechanism must have at least as much variety as what it is attempting to regulate (Ashby, 1956; Bar-Yam, 2004; Umpleby, 2008). An ability to generate new options is crucial if the environment changes beyond

already understood parameters. Requisite variety can be maintained by a learning network.

Missing from the preceding description of CAS are two features also commonly ascribed to CAS, i) the absence of central control, and ii) the idea that simple rules can explain system behaviour. These principles are explicit in Mitchell's (2009) definition of a complex adaptive system:

> *"A system in which large networks of components with no central control and simple rules of operation give rise to complex collective complex behaviour, sophisticated information processing and adaptation via learning or evolution"* (Mitchell, 2009, p. 13).

The notion of CAS functioning without central control is one of the common arguments raised against 'top-down' management of healthcare systems. The principle of simple rules is critical to this research, as it distinguishes a CAS explanation of change within a complex socio-technical setting from other organisational change paradigms.

A complex *adaptive* system is differentiated in that it learns from and adjusts to change. This interpretation can be placed within the *systems theory, complex systems theory,* and *cybernetics* foundational complexity science 'traditions' (Castellani & Gerrits, 2021). At times the healthcare QI literature borrows complexity terms from across all these contributing complexity disciplines, and a general sense of how the theories and tools fit together is helpful. Of the remaining two principal complexity thinking domains, *dynamical systems theory* is the domain of the mathematics and physics of complexity and has given rise to *chaos theory*, *non-linear systems theory* and *game theory*. The *artificial intelligence* tradition is the basis for *computational complexity theory, cellular automata* and computational techniques including agent-based modelling. These mathematics and computer science complexity fields play a valuable role in establishing CAS theory by providing many of the tools used to study, measure and understand complex systems.

2.3. Definition: Quality Improvement

Quality improvement can be used in a very general sense for any intentional change that improves one or more dimensions of quality. An all-encompassing definition for this general sense of QI is provided by Batalden and Davidoff (2007):

> *"the combined and unceasing efforts of everyone — healthcare professionals, patients and their families, researchers, payers, planners and educators — to make the changes that will lead to better patient outcomes (health), better system performance (care) and better professional development"* (Batalden & Davidoff, 2007, pg. 2).

This definition does not distinguish methods or tools, and offers hints that QI is as much an attitude as an intervention. Furthermore, it is holistic, involving 'everyone'. Similarly general in intent, Backhouse and Ogunlayi (2020) describe QI as:

> *"A systematic continuous approach that aims to solve problems in healthcare, improve service provision, and ultimately provide better outcomes for patients"* (Backhouse & Ogunlayi, 2020, pg.1).

Rather than specific methods or tools, it is the concept of systematic principles that is important for effective QI. Consistency, iterative testing of ideas, empowered staff and consumers, data-driven actions, and local adaptation are important healthcare QI principles. QI actions thus complement broader healthcare change strategies such as clinical audits and service evaluations (Backhouse & Ogunlayi, 2020).

'QI' is more commonly used in the healthcare literature to refer specifically to the practice, methods, and tools of quality management (Braithwaite, 2018; Dixon Woods & Martin, 2016). QI, when used in the narrower sense, is only one of several methodological approaches to achieving a general 'change for the better' in healthcare. This study encompasses both the general and specific meanings because both usages are intertwined throughout the literature, as shown in Section 3.2.

2.4. Further Relevant Definitions and Concepts

Researching complexity requires a clear definition of the type of complexity under consideration. Definitions and associated measures of complexity vary across physics, mathematics, biology, and computer science. Ambiguous definitions of complexity inhibit practical application and measurement (Adami, 2002). Fortunately, the various definitions tend to group along domain-specific lines. Biologists have concentrated on *structural complexity*, focusing on the number and variety of features that result from evolutionary processes. Physicists have developed an understanding of *process complexity*, differentiating chaotic, stable, and complex but statistically predictable

processes (Adami, 2002). A large number of complexity descriptions have been explored from these starting points. Mitchell (2009) provides a list that includes: *complexity as size; complexity as entropy*[2]*; complexity as algorithmic information content; complexity as logical depth; complexity of thermodynamic depth; complexity as computational capacity; statistical complexity; complexity as fractal dimension; complexity as degree of hierarchy.* Mitchell acknowledges the list is incomplete. The list is provided here only to emphasise how important it is to always qualify 'complexity' with the exact subtype under consideration.

Other types of complexity definition are potentially more applicable to socio-technical CAS. *Ontological complexity* has been used to describe 'real-world' complexity, distinguished from *cognitive complexity*, which encompasses the variety and subjective perception and description of complex phenomena (Rescher, 2019, as cited in Jackson, 2020). Cognitive complexity is not unlike the idea of *normative complexity*, which describes multiple conflicting values, perceptions, and habits amongst teams (Cribb, Entwistle & Mitchell (2021). Another important philosophical distinction introduced by Morin (2019) is the use of *restricted complexity* to refer to complexity that can be modelled and formalised. In contrast, *general complexity* cannot be meaningfully simplified enough to allow scientific analysis. In a socio-technical setting, general complexity is thus a similar concept to the *wicked problems* phenomenon coined by Rittel and Weber (1973). Wicked problems have such a high level of conflicting interpretations that they cannot be resolved objectively (Jackson, 2020; Morin, 2019; Rittel & Webber, 1973).

This study uses Wiesner and Ladyman's (2019) 'conditions for complexity', as an operational definition appropriate for the context of healthcare QI. These conditions describe the underlying *drivers* of complex phenomena: (i) numerosity of elements, (ii) numerosity of interactions, (iii) disorder, (iv) non-equilibrium and (v) feedback. Complex systems may also generate observable signature outcomes, described as 'products of complexity': (i) nonlinearity, (ii) self-organisation, (iii) robustness, (iv) adaptive behaviour, (v) modularity, and (vi) memory (Wiesner & Ladyman, 2019). Although not specific to CAS, the separation of complexity 'features' into a small set of generating conditions and observable effects facilitates consistent assessment of cases and cross-case

[2] The reference here is to Shannon entropy, the average information content a message contains

comparison. This approach makes it possible to capture at least some dimensions of complexity even with the limited narrative descriptions supplied in many cases.

CAS literature describes *Agents* as autonomous, bounded individuals or objects that employ adaptive strategies (Holland, 2006, 2012; Wilensky & Rand, 2015; Wilson & Kirman, 2021). This definition suggests that agents have a level of control over their actions, which is also described as *agency*. Appropriate consideration of the type of agency being described is necessary to meaningfully apply natural CAS explanations into the human socio-technical context. Agency is the organisation of behaviour where an agent directs its actions towards goals (Tomasello, 2022). Essential to a full consideration of agency is the recognition of graduated levels of agency. Tomasello posits that evolution has produced four levels of agency: (i) *Goal-Directed Agents* (using simple feedback control), (ii) *Intentional Agents* (with working memory), (iii) *Rational Agents* (with metacognition), and (iv) *Normative Agents* (with social and self-regulation). Dennett (2017) describes a very similar hierarchy of agency using the labels of *Darwinian, Skinnerian, Popperian* and *Gregorian creatures*. Darwinian creatures have hard-wired knowledge unable to be extended. Skinnerian creatures can learn via reinforcement feedback and Popperian creatures are able to form hypotheses. Finally, Gregorian creatures can construct and use thinking tools (Dennett, 2017). The typology proposed by Tomasello (2022) is perhaps the most useful for this research in that it incorporates the social organisation of agent behaviour (Tomasello, 2022).

Agents may interact in various ways, including competing or cooperating with one another. *Cooperation* describes agents who are prepared to pay an individual cost of some kind in return for another individual or group to benefit (Nowak, 2006) Applied to this study, cooperation is used to describe agents applying attention and effort (their personal cost) towards a shared improvement goal. Conversely, *rejection* describes an agent decision to stop cooperating with other agents. Groups of agents and their interactions can be thought of as an *agent network*, where the agents are represented as nodes in a graph and the links between nodes show the connections or interactions between agents (Mitchell, 2009). Network analysis, using the mathematics of graph theory, can then be used to evaluate the structural connections between agents (Barabási, 2016). In social network analysis, different measures of network centrality identify different subgroups of agents, the lengths of connected paths between agents, the number of connections agents have, and the vulnerability of the network to losing highly connected nodes. Most CAS

will be made up of many different networks of adaptive agents. In natural CAS, *coevolution* occurs when the successful adaptation of one agent or agent group is dependent on the adaptations of other agents or agent groups (Holland, 2012). This research retains the coevolution concept for socio-technical settings, in line with Holland's assertion of generalisable CAS behaviour. Coevolution serves as a useful description of change within an agent network that requires an overlapping or external agent network to also change.

Emergence describes phenomena that arise from the structure and interactions of the system, occurring at a different scale than the component parts of the system. Emergent phenomena are thus macro-level outcomes dependent on the micro-level interactions. For example, a pattern of traffic arising from the interactions of many vehicles (Carmichael & Hadžikadić, 2019). Similarly, 'flight' emerges from a correctly assembled and operated set of airplane parts, and 'quality' can be considered an emergent outcome of operational processes.

A system *environment* describes the biological or socio-technical setting within which agents and agent networks interact. In natural CAS, this is the natural ecological habitat of the agents, for example, a forest ecosystem. In this research, the environment describes the socio-technical setting within which networks of agents function. For example, a workplace, an organisation, or a city. The boundaries between individual agents, and between agents and their wider environment, are inherently subjective and dynamic (Holland, 2012; Nair & Reed-Tschocas, 2019).

Complex systems can exhibit *non-linear* outcomes. Processes where cause and effect are not directly proportional on the same scale, for example a power law or exponential process, are described as being non-linear (Wiesner & Ladyman, 2019). Non-linear is also used to describe a system where the whole is not equal to the sum of the parts. In addition to power law effects as noted, non-linearity may be caused by discontinuous step changes in system states or 'tipping points' that trigger cascading effects (Mitchell, 2009).

Schema is used to describe the Agent's internal mental models of their environment and their internal rules for action (Dooley, 1997). Schema thus function as the agent's interpretation of the system and appropriate actions to take. *Shared schema* is posited in this research to mean a level of congruence or alignment in mental models between agents (Cronin & Weingart, 2007). Shared schema are not identical across agents but are in

agreement relative to a goal. *System signals* send information about the system structure, behaviour and interactions as feedback to update agent schema (Holland, 2012). This study presents the overall process of responding to environmental (system) signals, reviewing internal mental models, and then taking action as a form of *problem solving*. At a more granular level of description, problem solving involves selecting *elements* (that can be manipulated and *operators* (actions that can be taken) within a set of possible actions (the *problem space*), until the goal state is reached (Anderson, 1993, 2005; Newell & Simon 1972).

Complex systems create *uncertainty* in multiple ways. There may be so many system parts and interactions that the system cannot be adequately perceived or understood. Cause and effect within agent actions may not be able to be meaningfully linked. Multiple system outcomes may be possible for any given action. Incomplete knowledge of the current state of the system and an inability to be certain of cause and effect are forms of *epistemic uncertainty* (Hüllermeier & Waegeman, 2021). Epistemic uncertainty is distinguished from *aleatoric* uncertainty, which is a product of random chance or natural variation within a process. Importantly, epistemic uncertainty due to incomplete knowledge can in principle be reduced. Both forms of uncertainty can be assessed and quantified using probability theory (Fenton & Neil, 2019; Hüllermeier, & Waegeman, 2021).

Chapter 3 Literature Review

3.1. Introduction

Interpreting healthcare QI as CAS involves the intersection of three overlapping knowledge domains: (i) healthcare; (ii) quality improvement, and (iii) complex adaptive systems. Each of these domains is reviewed to establish the main threads of current understanding, ongoing questions, and potential research gaps. An overall picture is compiled showing relevant commonalities and distinctions as they affect the topic. In other words, what is known about this intersection and what remains to be understood?

Section 3.2 reviews QI as described and researched in the healthcare QI literature. Important healthcare improvement methods distinct from QI, such as *implementation science*, are discussed. The extent to which these methods overlap with QI or are fundamentally different is considered. Section 3.3 reviews CAS principles as they have been set out in a range of theoretical and applied sources. This groundwork is necessary to establish the consistency (or boundaries) of CAS as an analytical lens for healthcare QI. Section 3.4 analyses the relationship between quality management thinking and complexity, independent of the healthcare QI context. Relevant management and organisational theories are also discussed in this section. This section asks if healthcare explorations of complexity are substantially different to other domains. To understand the state of modelling as a research method for QI, a review of recent modelling approaches to QI is presented in section 3.5. A comprehensive scoping review process was used to identify secondary data cases, and this process is described in section 3.6. Note that relevant theoretical perspectives also identified in the scoping review have been integrated into the syntheses above. Bringing all the reviews together, section 3.7 recaps the essential themes of the literature review, and finally the research gap is described in section 3.8.

3.2. Synthesis 1: Healthcare QI and Complexity

3.2.1. The Ecosystem of Healthcare Change Strategies

Healthcare has evolved multiple strategies for change beyond traditional quality management. These approaches have a basis in the unique nature of health and healthcare delivery. Additionally, clinicians have not limited themselves to quality management and drawn inspiration from other knowledge domains, such as social policy and safety engineering. A persistent challenge in healthcare delivery has been reducing the gap

between currently understood best practice for a given condition, and the delivery of care related to that condition in practice. Implementation of best clinical practice is not the same thing as researching best clinical practice (Bauer et al., 2015; Braithwaite et al., 2018). Addressing this gap, *Implementation Science* is conceived as a mirror complement to the clinical research and evidence-based medicine (EBM) components of healthcare. How can the latest research be shared widely and incorporated into practice? It is a trap to imagine research dissemination as a linear pipeline from scientific research, standardised intervention designs and implementation into practice. This 'pipeline' view fails to account for the substantial variation of local contexts, and the ability of local actors to contribute or respond. (Hawe, 2015). Implementation science is defined as *the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice* (Eccles & Mittman, 2006, as cited in Bauer et al., 2015, pg. 3). Thus, implementation science applies evidenced-based thinking and practice to the successful local adaptation and uptake of clinical evidence.

Whereas implementation science focuses on intervention to bring best practice to the point of care, *Learning Health Systems* (LHS), take a macro-level approach to turning knowledge into new practice at scale. LHS emphasise health informatics and computable knowledge including electronic health records. By integrating data and data science techniques into clinical practice and culture, LHS aspire to do more than distribute or share knowledge. *New* knowledge is made available from systemic learning cycles where data becomes knowledge, knowledge becomes practice and practice becomes data (Adler-Milstein, Nong, & Freidman, 2018; McLachlan et al., 2018). LHS thinking explicitly recognises the knowledge management and knowledge work involved in healthcare delivery and leverages advances in computer and data science alongside clinical networks and quality improvement activity (Seid, Hartley, & Margolis, 2021).

Another critical dimension of health care quality is patient safety. The field of *Resilient healthcare systems*, or *resilience engineering*, draws on the pioneering work of Hollnagel, Wears, and Braithwaite (2015). Resilient healthcare systems are based on a concept of safety called *Safety-II*. Safety-II challenges linear thinking and proposes an approach to patient safety better able to cope with complex environments. In Safety-I thinking, the emphasis is on identifying and preventing the root causes of harmful outcomes. Safety-II thinking regards safety is an emergent outcome of many interacting processes, thus requiring a different approach. Concentrating on what is *working well*

helps build up the understanding of the system, including necessary and sufficient conditions for safety. Additionally, human actors can be viewed as adaptable and resilient change agents to proactively manage safety, rather than being intrinsically fallible. Safety-II is not proposed as a mutually exclusive alternative to Safety-I — it is a complement to conventional incident investigation and root cause analysis (Hollnagel, Wears, & Braithwaite, 2015).

Deep consideration of effective social policy has influenced another important healthcare improvement strategy. There is an obvious connection between social policy and population and public health programmes, for example, encouraging child immunisation or healthy diet promotion. *Service Improvement* takes the scientific evaluation of social policy changes and applies the tools to healthcare service delivery, especially public health (Campbell, 1971; Funnell, & Rogers, 2011). The two central ideas of service improvement are Program theory and Evaluation. Program theory is a holistic approach to change that specifies careful articulation of the hypotheses behind any planned change. Exponents of program theory identify the theory of causation, theory of change, and theory of action, that underpin an intervention (Aragon, 2017). The applied theories (hypotheses) behind an intervention are represented via logic models, which provide a 'map' of change and an analysis tool to support evaluation (Funnell & Rogers, 2011). Formal evaluation is critical to service improvement, given the scientific goal of supporting hypotheses with evidence, as well as more practical requirements such as securing funding.

The preceding paragraphs reveal how different healthcare-specific contexts — the implementation of clinical research, managing patient safety, and evaluating social change — sit alongside the specific tools of quality management in healthcare. Despite the differences in origin, collectively they form a kind of ecosystem with specific niches and the different change approaches are regularly used in a complementary fashion (Aragon, 2017; Braithwaite, Marks, & Taylor, 2014.; Britto et al., 2018b; Rogers, 2008). Having surveyed healthcare change strategies at the general level of 'improvement', the context is established to now explore QI methods in detail.

Quality improvement, used in the narrow sense of a specific set of methods and tools, is widely known as 'Improvement Science', particularly within healthcare. Deming's (2000) 'System of Profound Knowledge' serves as a capstone structure for

improvement science, drawing on a body of knowledge stretching back to Shewhart and Deming's early twentieth century work in statistically-driven industrial process control (Lemire, Christie & Inkelas, 2017; Perla, Provost & Parry, 2013). The system of profound knowledge connects four pillars of improvement: (i) appreciation for a system; (ii) understanding variation; (iii) theory of knowledge, and (iv) psychology (Deming, 2000). Throughout his long career, Deming experienced and well understood many of the 'system' notions introduced in section 3.3. Organisations are made up of people, and quality will not be meaningfully improved without understanding how humans develop and interpret knowledge, and also how they interact in socio-technical contexts. Appreciation for a system also recognises relationships between system parts alongside the actual structural elements. Today, Deming and his teacher Shewhart are best known for statistical process control (SPC), applying quantitative statistical techniques to establish the levels of variation in processes. SPC distinguishes common (random) cause variation from special cause variation, where an underlying parameter has changed (Deming, 1986). Potentially, the deeper system thinking that was undoubtedly part of Deming's legacy has been somewhat overlooked compared to the impact that SPC has had on quality management. Deming developed *plan-do-check-act* cycles as a scientific learning process where a potential change is identified, tested, reviewed and then revised in an iterative manner (Beckford, 2002). The parallel with scientific method is intentional and contributes to the wide acceptance of the approach in which the QI professional: (i) proposes a hypothesis; (ii) tests that hypothesis; and (iii) either accepts, rejects, or modifies the hypothesis (Deming, 1986). Over time Deming evolved his cycle to become *plan-do-study-act* (PDSA). 'Study' better represents the deeper analysis required to evaluate a test, as opposed to simply checking against set testing criteria (Moen & Norman, 2016).

PDSA cycles are an essential and ubiquitous element of QI in healthcare today, either in their original form or extended into healthcare-specific change methods such as the *Model for Improvement,* developed by The Institute for Healthcare Improvement. This version augments PDSA cycles by making explicit the goals, selection, and measurement of change. Essential to the PDSA method is predicting the outcomes of a test, in direct analogy to scientific method. PDSA cycles mix inductive and deductive logic modes and represent 'learning by doing', which was also an important aspect of Deming's

philosophy of acquiring knowledge (Deming, 1986; Institute for Healthcare Improvement, 2023; Lemire, Christie & Inkelas, 2017).

Another healthcare adaptation of QI is the idea of *quality improvement collaboratives* (QIC). As the name implies, QIC are focussed efforts at QI by groups of healthcare professionals, usually taking a specific issue, and using the collaboratives to build relationships, share knowledge and generate solution ideas (Schouten et el., 2008). The concept of iterative PDSA cycles to test and learn remains important. The production sector is very different, but it is notable that QIC are similar to Kaizen or worker quality circles, both common notions in manufacturing quality thinking (Imai, 2012). The linking ingredients are the professional autonomy to define a problem to work on and then developing/sharing peer knowledge to try and resolve problems.

Lean, SixSigma, and Lean SixSigma are the other major QI methods that have been widely applied to healthcare service provision. SixSigma is a modern codification of statistical process control, complete with a phased project method to test and deliver change. Arguably, Lean has had the most influence in healthcare quality thinking (D'Andreamatteo, Ianni, Lega & Sargiacomo, 2015). In 1996 Womack and Jones published their book *"Lean Thinking: banish waste and create wealth in your corporation".* This book introduced the label "Lean Thinking", positioning the Toyota Production System and the Toyota Way as a generalised management system applicable at an enterprise level and in industry sectors beyond manufacturing (Womack & Jones, 2003). The five fundamental components of Lean Thinking are: (i) value — define value to your customer; (ii) value stream — identify the steps in your process that create value; (iii) flow — refine your process to link value adding steps and eliminate waste; (iv) pull — the consumer triggers the production/value creation, and (v) perfection — seek perfection via continuous improvement.

Moving from Lean Manufacturing to Lean Thinking at this higher level of abstraction introduces ideas that are not out of place in healthcare. In particular, the concepts of flow (connecting sequences of value-adding steps), minimising unwanted variation and eliminating waste are promoted (D'Andreamatteo et al., 2015). Lean thinking, along with the other industrial QI methods and tools outlined here, has been applied to healthcare quality endeavours for decades. The effectiveness of this strategy

remains an active topic of debate. A question recurs in the healthcare improvement literature – is QI improving health outcomes?

3.2.2. Is Healthcare QI Effective?

Reliability and quality problems within healthcare exist at a level above that accepted in other high-risk sectors (Braithwaite, 2018; Dixon-Woods & Martin, 2016). One reason for this situation may be a lack of adherence to the critical principles of improvement methods. Also recognised is the need to shift from isolated, small-scale projects to better-coordinated programmes of activity that can achieve success on a significantly greater scale. Essentially, there is a constant loss of learning resulting from the failure to iteratively build on and share QI lessons (Dixon-Woods & Martin, 2016). However, a prerequisite to better coordination and sharing of QI effort is standardised QI reporting, something still regarded as inadequate by many researchers (Dixon-Woods & Martin, 2016; Thompson et al., 2016). Improvements in reporting consistency have been achieved through frameworks such as the *Consolidated Framework for Implementation Research* (CFIR) developed by Damschroder et al. (2009). CFIR provides a typology of domains and characteristics for conducting research into healthcare interventions (Damschroder et al., 2009). CFIR is complemented by the SQUIRE 2.0 Guidelines (Standards for QUality Improvement Reporting Excellence), which detail the specific documentation requirements to ensure ease of sharing and interpretation of QI research findings (Ogrinc et al., 2015). These developments are important but will take time to become widely established as standards 'in use'. As part of the design for this research, database search queries were conducted for SQUIRE-compliant QI research articles that also reflected on complexity in some manner. The source pool of potential articles for review was very limited when restricted to this extent.

As noted, QIC are one improvement approach that can be said to have developed organically within healthcare's unique cultural and organisational structures. We might therefore expect that QIC are able to navigate structural barriers to change. The evidence for this is positive, but limited (Schouten et al., 2008). There are methodological concerns with how outcomes were reported, and, as with other improvement strategies, a lack of detailed, standardised reporting on necessary elements of the organisational context to facilitate wider uptake (Schouten et al., 2008).

Another concern expressed in the healthcare QI literature is the importance of fidelity to the requirements of a given change method. It is difficult to conceive of more conceptually straightforward QI framework than *Plan-Do-Study-Act* (PDSA) cycles. Despite the sound theoretical basis and its structural simplicity, implementation challenges are common (Reed & Card, 2016). Research summaries reveal poor compliance with the method. Alarmingly, fewer than 20% of articles reviewed adequately document the iteration process, which is essential to the learning outcomes and improvement refinement. The need for standardised reporting as part of the solution to achieving consistency is repeated and there are calls for increased scientific rigour in the application of this method (Taylor et al., 2013, p.8). If healthcare organisations struggle with PDSA cycles, it is unsurprising that they might find it difficult to achieve convincing results using QI methods requiring more investment from an organisation such as Lean, Six Sigma, and QIC.

The effectiveness of Lean in healthcare has been extensively studied, with several socio-technical and cultural obstacles noted. The healthcare-specific cultural barriers to Lean are multiple. Lean (and other QI methodologies of non-healthcare origin) can disrupt the established hierarchies within and across professional workforce groups (Brandão de Souza & Pidd, 2011; Waring & Bishop, 2010) Without full understanding of the evolution and strength of the intricate relationships that exist in multi-disciplinary teams such as clinical microsystems, any non-context sensitive QI intervention is likely to meet resistance (Waring & Bishop, 2010). The notion of *value* that is central to Lean is also problematic in healthcare. 'Whose value is being created?' is the critical question. Unless clinical, operational, and experiential dimensions of value are precisely defined and agreed upon by everyone involved in an improvement initiative, the motives for using Lean may be questioned, and assessment of success (i.e. achieving more of the desired value) will be open to dispute (Young & McClean, 2008).

Agreeing an acceptable standard of evidence for QI is another factor. There is limited evidence for the effectiveness of QI from controlled, repeatable experimental trials — which are the established norm in the assessment of clinical effectiveness (Walshe, 2007). Concerns over the evidence for Lean's effectiveness in healthcare relate to the over-emphasis on case studies in the literature; however, several commentators (e.g. Kaplan et al., 2011; Walshe, 2007), point out that quality improvement in healthcare is a *"complex social intervention"* (Walshe, 2007, p. 57), and therefore not well suited to experimental

research design methods. Walshe does not advocate for the abandonment of experimental method altogether, but he argues that experimental designs must be complimented by a theoretical approach, as is the norm in other social science research such as education and justice (Walshe, 2007). Other researchers echo this call for more theory-based research into improvement (Davidoff, Dixon-Woods, Leviton &Michie, 2015; Greenhalgh & Papoutsi, 2018; Khan et al., 2018).

An obstacle to experimental methods is the ability to isolate and control relevant experimental parameters. The idea of *context* becomes critical, and the principle that in complex socio-technical settings, the intervention and context work together. This is of course the key principle of implementation science (Bauer et al., 2015). Several models of contextual factors have been developed and applied to examine contextual factors associated with any QI activity. The *model for understanding success in quality* (MUSIQ) developed by Kaplan, Provost, Froehle, and Margolis (2011) has received some attention and been further explored in independent research (Barson, Doolan-Noble, Gray, & Gauld, 2017; Griffin et al., 2017; Kringos et al., 2015; Michel et al., 2016). The MUSIQ contextual model is hierarchical and considers key activities at the external environment, and multiple organisational, layers. QI culture, for example, is assessed at organisational, project leadership and participating staff levels. The most frequently identified factors influencing hospital QI strategies are: (i) the clinical microsystem; (ii) QI support; and (iii) the capacity of the QI team (Kringos et al., 2015).

Contextual models like MUSIQ have limitations. For example, how system actors communicate is not clear. Static taxonomies of factors to measure and control are helpful, but incomplete. The highly dynamic interactions present in a clinical microsystem are not sufficiently considered. In response to these criticisms, MUSIQ continues to evolve (Reed, Kaplan, & Ismail, 2018). Interactions between contextual factors are more explicit, and the non-linear relationship between interventions and outcomes is acknowledged. Adequately supported, iterative, microsystem-level QI activity is used as the vehicle to learn and adapt to contextual uncertainty. A criticism that remains is that the interactions between agents (rather than the structural elements) are positioned at a very high level, as ‘cultural’ factors. The potential barriers to healthcare QI effectiveness considered so far are now summarised in Table 3.1, with the impact of these barriers also noted:

Table 3.1: Barriers to QI Effectiveness

Potential Barriers to QI Effectiveness	**Impacts**
Fragmented or disconnected QI initiatives	Lack of critical mass for transformational change.
Inconsistent reporting standards	Difficulty sharing and learning. Duplication of efforts.
Lack of fidelity to QI methods	Sub-optimal results. Continuing ambiguity as to whether the methods are effective.
Cultural fit with healthcare	Resistance to full participation.
Dynamic and highly varied contexts	Inability to replicate successful interventions. Difficult to assess impact via randomised control trials.

The barriers listed in Table 3.1 reflect a progression in difficulty from technical considerations through to more systemic challenges. Achieving transformation at large scale requires coordinated effort, whereas QI initiatives are predominantly locally driven and not easily shared. Consistent reporting standards do exist but will take time to become widespread. Analysis of QI implementations reveals that the important tenets of methods are not always followed, which will naturally undermine effectiveness. Questions of cultural fit for traditional QI have been raised, even once the superficial 'not invented here' and 'people, not widgets' instinctive reactions to industrial QI are overcome. The rich human interactions required in health may simply not be suited to prescriptive, linear methods. Finally, the highly variable and dynamic settings require a thorough understanding of the unique context for *each* intervention. Taken together, these themes suggest there are very considerable barriers to effective healthcare QI. It is possible that more patience is required with current strategies to see beneficial outcomes. It will take a long time, and adequate investment, to build a critical mass of healthcare improvement

knowledge and competency (Mondoux & Shojania, 2019). However, other researchers have suggested that what is actually required is a fundamentally different approach, one that can adequately function in complex settings.

3.2.3. Healthcare Interpreted as a CAS

CAS ideas and language resonate strongly with healthcare researchers (Braithwaite et al., 2018; Chandler et al., 2015; Khan et al., 2018; Lorden et al., 2014; Thompson et al., 2016; Turner & Baker, 2019). Conceptualising healthcare as a complex system is often interpreted as a rejection of reductionist solutions that ignore interconnected system relationships and the use of deterministic, "machine" metaphors for the system (Mahajan, 2017; Young et al., 2020). The implication is that something very different is taking place within CAS. In a CAS model of healthcare QI, system boundaries are no longer rigid and interactions between agents in the system are entangled and unpredictable. Agents apply internalised rules and self-organize their collective behaviour. Successful change will require coevolution with intersecting agent networks. Outcomes of interest within a healthcare QI CAS are emergent, difficult to predict, and non-linear. Notwithstanding these difficulties, meaningful patterns of change may be observable, even if they do not yield to precise analysis. Plsek and Greenhalgh (2001, pg. 627) go as far as stating that: "*ultimately, the only way to know exactly what a complex system will do is to observe it".*

Scepticism towards the applicability of CAS principles within human sociotechnical contexts is rare in the literature but does exist and raises valid concerns. Yawson (2013) describes the wide gap that exists between theoretical papers and practice within human resource development research. A prerequisite of teaching the background systems theory is missing and there are few attempts to purposefully apply complexity theories as an alternative to linear frameworks (Yawson, 2013). Within the healthcare as CAS discourse, questions over the role of human cognition and system design within CAS have been posited. Paley (2010), and Paley and Eva (2011) have argued that cognition and self-awareness result in a very different process to natural CAS. The position taken is that complex goal-oriented activity cannot *literally* be a CAS because CAS by definition lack a system-level design and agent intentionality (Paley & Eva 2011). The point is somewhat technical, relating to the validity of explanations for how CAS function, and Paley and Eva do not dismiss the usefulness of CAS as an analytical lens. A further tension may exist between QI and CAS in terms of fundamental paradigms. QI focuses on reducing gaps between actual and desired quality outcomes;

promoting the ‘best’ way to achieve something inevitably reduces variety and eventually leads to inertia. This tendency is in direct conflict with the CAS principle of increasing variety so that a point of requisite (sufficient) variety is reached (Ashby, 1956; Junghans, 2017; Bar-Yam, 2004).

Strategies, conceptual frameworks, and tools for tackling complex problems have been developed, such as those listed in Table 3.2 below. However, the problems in healthcare delivery have been shown to be persistent (Braithwaite, 2018; Dixon-Woods & Martin; 2016). The applied CAS research field still contains gaps and exploring these gaps thoroughly may lead to further innovation and more effective methods.

Some clues as to what might be missing can be seen when we consider existing complexity management frameworks as a group. While not intended to be exhaustive, Table 3.2 lists six well-known examples and notes the high-level emphasis of each framework. Inclusion in this list was based on maturity of the approach and observation of active use supporting QI research and practice (Anderson et al., 2015; Augustsson et al., 2020; Bäcklander, 2019; Kringos et al., 2015; Perla et al., 2013; Snowden & Roncati, 2021). But with the exception of MUSIQ, these are not healthcare-specific frameworks. There are thematic commonalities, most notably the emphasis on situational awareness and specific leadership/management behaviours. QI Practitioners interested in working with complexity principles have multiple options for guidance, and increased variation along these same themes would be difficult to justify.

Table 3.2: Frameworks for Complexity-Informed Quality Improvement

Framework	**Short Description**	**Thematic Emphasis**
Soft Systems Methodology (Checkland, 2000)	An approach to analysing and responding to complex problems. *Soft* systems assume the presence of complex interactions and multiple valid perspectives of problems and solutions, in contrast to clear or stable (*hard*) systems. A cyclic learning "system" is used to propose, explore and test potential solutions from multiple perspectives.	Learning cycles Multiple stakeholder perspectives
System of Profound Knowledge (Deming, 1994)	Contains four core elements: 1. Appreciation for a system (knowledge of interdependent parts and their interactions). 2. Theory of knowledge (what we know and how we can acquire new knowledge). 3. Psychology (understanding behaviour). 4. Understanding variation (common and special cause variation).	Learning cycles Interdependence Human behaviour

Framework	**Short Description**	**Thematic Emphasis**
Cynefin (Kurtz and Snowden, 2003)	A sense-making, situational awareness and decision support framework. Cynefin proposes that systems dynamically move between ordered and unordered domains, further broken down into *clear*, *complicated*, *complex* and *chaotic* system states (or problem scenarios). Leadership, decision-making and intervention actions differ based on the current system state and the level of knowledge. Optimal management responses differ across different system states.	Situational awareness Sense-making Leadership skills and behaviour Decision making
Complexity Leadership Theory (Uhl-Bien, Marion, and McKelvey, 2007)	Leadership skills suited to knowledge-oriented activity rather than production focussed. *Adaptive* outcomes – learning, innovation, adaptability. *Entangled* leadership roles – adaptive, administrative and enabling leadership functions. Emphasizes the dynamic tension between administrative, bureaucratic functions and emergent, informal interactions.	Situational awareness Leadership skills and behaviour Decision making

Framework	**Short Description**	**Thematic Emphasis**
Adaptive Leadership (Heifetz, Heifetz, Grashow and Linsky, 2009)	Proposes leadership requirements based on situational complexity. Adaptive leadership fosters skills required to navigate *technical* and *adaptive* scenarios as appropriate. Simple problems require technical responses, complicated problems require technical and adaptive responses, and complex problems require adaptive responses. Balance between top-down, standardised procedure responses to clearly defined issues and bottom-up, staff-led learning to understand complex issues and build solutions.	Situational awareness Leadership skills and behaviour Decision making
Model for Understanding Success in Quality (MUSIQ) (Kaplan, Provost, Froehle and Margolis, 2011)	Contextual factors model with recursive macro, meso and micro levels. An assessment tool to examine enabling competencies for change. Encompasses leadership, quality management teams, and clinical microsystem teams that support and implement change.	Context and enabling factors Leadership hierarchy

The complexity frameworks introduced above tend to emphasise structure and leadership. Is this emphasis appropriate? Potentially, systems thinking may be vulnerable to hierarchical and complexity biases; we may unconsciously introduce hierarchy and complexity into our systems thinking even when this is not justified (Carmichael and Hadžikadić, 2019). In contrast, explanations of natural CAS behaviour are usually associated with a small set of simple agent rules and avoid hierarchical agent structure (Holland, 2006). The frameworks summarised in Table 3.2 access CAS principles *indirectly* through the filters of our socio-technical norms, such as structures, leadership hierarchy and decision-making. *Directly* harnessing CAS phenomena, free from hierarchical and complexity biases, is a valid and still under-explored alternative strategy (Carmichael and Hadžikadić 2019).

Despite the availability of such complexity-informed improvement tools as introduced in Table 3.2, concerns remain over the effectiveness of healthcare QI specifically under complex conditions (Braithwaite et al., 2018; Dixon-Woods and Martin 2016). The inference being made here is not that the current strategies or tools are ineffective, but that CAS thinking may still have more insights to offer for change within complex socio-technical systems. Given the continued pressure on healthcare systems and the potential impact of effective QI, it is important to explore all options comprehensively. Specifically, the way change occurs within sophisticated *human* CAS lacks explanatory mechanisms. What replaces the biological and evolutionary processes within natural CAS? More complete theory will allow practitioners to 'join the dots' across many unique contexts and extract the essential principles from disparate complexity-informed tools.

It is encouraging to note the extent to which researchers are using complexity ideas, including CAS principles, to complement and extend existing knowledge and frameworks. The use of network science is one example of this, where researchers are studying the interactions between system agents as a fundamental element of QI (Britto et al., 2018b; Cunningham et al, 2012; Kitson et al., 2018; Levula, Chung, Young & White, 2013; Manukyan, Epstein & Horbar, 2013). Basic QI, including PDSA cycles, remains important in these examples, now enhanced with better contextual awareness, communication, and learning. Learning is another critical theme within healthcare as

CAS thinking. In the presence of uncertainty, we need to work out (learn) what is happening and how best to respond. C*ollaborative* learning is essential to resilient healthcare systems (Haraldseid-Driftland et al., 2022). Furthermore, learning is not an isolated or external activity, but deeply embedded into normal clinical and administrative activity. Learning is also integrated into any QI actions. At the level of PDSA cycles, double-loop learning — revising the hypothesis in the face of new information — is critical to attempting change in dynamic and uncertain contexts (Aragon, 2017).

The 'ecosystem' of healthcare improvement is merging and creating hybrid methods. Healthcare-specific, CAS-informed QI frameworks are being developed. The framework for *Successful Healthcare Improvement From Translating Evidence in complex systems* (SHIFT-Evidence) represents the next generation of contextual factor models. SHIFT-Evidence includes three general principles for implementing evidence-based practice in complex systems, and twelve 'simple rules'. The three very general principles identified are: (i) *Act scientifically and pragmatically*; (ii) *Embrace complexity*, and (iii) *Engage and empower* (Reed, Howe, Doyle & Bell, 2018). Whilst these principles are admittedly very high-level, the key ideas of empowered agency, learning from doing, and accepting uncertainty are conveyed. Associated rules show how the principles can be enacted, but these also remain very high level. As representative examples: *Facilitate dialog; Build a culture of willingness to learn and freedom to act; Understand practices and process of care; Seek political, strategic and financial alignment* (Reed, Howe, Doyle & Bell, 2018, pg.10-11). A similar but slightly more prescriptive set of 'simple rules' to enhance success in large-system health care transformation is proposed by Best et al., (2012): (i) blend designated and distributed leadership;(ii) establish feedback loops; (iii) attend to history (i.e. recognise path dependence); (iv) engage clinicians, and (v) include patients and families. It is very important to note that the 'simple' rules described at this level of abstraction are not the same thing as the CAS concept of 'simple' rules contained within agent's internal schema. Braithwaite (2018) provides a similarly high-level list of "Twenty complexity-oriented enablers and insights", that includes such items as *look for behavioural patterns in the system and listen to the language people use; Beware excessively causal logic; Use multimethod research and improvement techniques* (Braithwaite, 2018, pg. 3). Such guidance from very experienced experts is obviously useful for improvement practitioners to consider and extremely important. The challenges

for practitioners might by working out how to follow the guidance, and also accepting that many of the suggestions can only be fully addressed by executive management.

The literature on Healthcare QI and complexity is extensive and is an active field of research. Clearly, researchers and practitioners are motivated to improve healthcare QI outcomes and still seeking effective methods relevant to their circumstances. Across this range of research, four important and consistent themes can be discerned.

First, there is no single change strategy, but multiple approaches that overlap and are often employed together. The impression is one of a rope braided together for greater strength than the individual strands. QI, in the narrow sense, is usually incorporated in all the change strategies in the form of systematic learning cycles. The limitations of a deterministic systems model are put forward by some researchers, but the literature is contradictory in terms of what this now means for practicing QI in healthcare. On the one hand, the CAS paradigm implies new methods are required. In contrast, other researchers are simply adopting complexity concepts to complement and enhance quite conventional improvement initiatives (Barach & Kleinman 2018; Britto et al., 2018b, Reed et al., 2018).

Second, something makes QI in healthcare difficult. The specific cultures within healthcare may play a role, but complexity conditions are also a factor — high numerosity of agents and interactions, feedback, lack of equilibrium and non-linearity between intervention and outcomes. The need for local adaptation of ideas is regularly emphasised. Braithwaite (2018) elaborates on the paradigm tension between resilient and emergent systems, and the ineffectiveness of ‘top down’ change. Healthcare systems are naturally resilient in the sense of buffering against changes that do not make sense to frontline clinicians delivering care. For Braithwaite, meaningful change is always emergent and local. A technical point noted here is that this ‘adaptation’ is frequently used in the sense of ‘modification’ of an idea to fit local contexts, rather than the larger process of taking action based on environmental signals. Another barrier to healthcare QI may be the very high *Normative Complexity* present in healthcare. Normative complexity refers to the inbuilt beliefs and values as to ‘what matters’, and how varied this can be across multiple professional and stakeholder groups within healthcare (Cribb, Entwistle & Mitchell, 2021).

Third, references to CAS in the healthcare QI literature are often limited to isolated features, and wrapped in superficial descriptions of complexity. The selective use of different complexity language and principles is the main finding of multiple comprehensive reviews of complexity in healthcare and social science literature (Chandler, et al., 2015; Thompson et al., 2016; Turner & Baker, 2019). This makes the literature hard to unpick. Ongoing refinement of theory-driven approaches, rather than isolated use of complexity concepts, is required (Greenhalgh & Papoutsi, 2018; Khan et al., 2018; Thompson et al., 2016).

Fourth, and finally, Carmichael & Hadžikadić's (2019) points regarding unintentional hierarchy and complexity bias in CAS thinking seems to be well made. Contemporary, complexity-aware QI methods proposed such as the SHIFT-Evidence framework still respond to complexity with hierarchical models and advice for leadership. The CAS principle of self-organisation within healthcare socio-technical systems is insufficiently developed beyond the need to adapt solutions in the modification sense noted previously. Researchers can point out that leadership must be practiced at macro, meso, and micro system levels, but healthcare delivery and healthcare QI still takes place within institutional structures that work against genuine self-organisation.

3.3. Synthesis 2: Complex Adaptive Systems

How does a complex adaptive system 'adapt'? Perhaps the clearest formal articulation of CAS function has been provided by John Holland (1992, 2006, 2012). Holland defines four essential CAS attributes: (i) *Massive Parallelism,* where large numbers of agents interact simultaneously by sending and receiving signals; (ii) *Conditional action*, where agent actions are conditional on signals and rules available to the agent; (iii) *Modularity*, where activities can be combined to form larger "subroutines", and (iv) *Adaptation and evolution,* where the agents change over time, via adaptations that improve a situation (rather than random change). Adaptation requires *credit assignment* (weighting and rewarding individual signals and rules within the parallel activity) and *rule discovery* (when the agent's existing rules are incomplete or inadequate). A formal algorithmic process, a *classifier* system, is used to simulate the micro agent behaviour of signal detection, signal evaluation, rule evaluation, rule generation and rule selection (Holland 2006, 2012; Urbanowicz & Moore 2009). Through his use of genetic algorithms, Holland created generalisable agents with in-built rules. These agents are also capable of creating new rules by various methods for recombination and then selecting which rule to apply

in response to a signal. This process is the algorithmic equivalent of genetic variety and selection — an explicit mimicry of evolution by natural selection. The emphasis is on signal interactions, the boundaries between agents, and the minimal set of rules that give rise to adaptive behaviour. The concept of agency is restricted to a basic level in the sense of *self-directed behaviour towards a goal*. This very rudimentary level of agency might be acceptable for computer generated agents or groups of social insects and animals, but it is unlikely to persuade sceptics that human socio-technical systems are genuinely CAS. Satisfactorily expanding or replacing this classifier "adaptive" kernel presents the biggest challenge in proposing that CAS principles are appropriate for nuanced, sophisticated human behaviour.

Conceptualising improvement as a problem-solving process provides a bridge between the rule-based behaviour of non-self-aware agents within a CAS and the goal-oriented activity of human agents. Adaptation directed towards improvement implies problem solving to some degree, which may range from trivial to extremely difficult. 'Problem' is used here without any value judgement; it is the stimulus for action of some kind to achieve a goal. Formal problem-solving strategies, such as *means-end analysis* or the *goals-assumptions-elements-operators* model offer a practical way of introducing purpose, awareness, and human agency into the agent adaptation process (Anderson 2005; Cronin & Weingart 2007).

Within natural CAS, inherited and learned behavioural rules, variation, genetic mutation, phylogenetic drift, and natural selection drive variation and selection (Holland 2006, 2012; Kauffman 2000; Maturana & Varela, 1987). Problem-solving processes provide an analogous learning and adaptation function in human endeavours (Anderson 2005; Doolittle 2014; Newell & Simon 1972). *Cooperation* amongst agents is required to achieve scale beyond individual agents. Cooperative problem solving establishes, tests, and evolves the mental models or schema of participating agents (Cronin & Weingart 2007; Dooley 1997; Mohammed et al. 2010). Problem solving requires knowledge which, if not initially present in the agent network, must then be acquired and learned.

Are there further connections between natural, engineered, and social systems? An ecosystems approach suggests six system organising principles that generate CAS features: (i) the system is constituted relationally; (ii) the system has adaptive capacities; (iii) the system has dynamic processes; (iv) the system is radically open; (v) the system

is contextually determined; and (vi) novel qualities emerge through complex causality (Preiser et al., 2018). The relational, adaptive, dynamic and emergent principles have already been outlined; the remaining two principles add important emphasis. General Systems theorists make careful distinctions between open and closed systems. An open system exchanges energy, information or materials with its wider environment (Ashby, 1956; Buckley, 2017). In this view, CAS are open systems with highly dynamic boundaries and subjective boundaries as these depend on agent context. Identifying the 'system' and defining its internal and external boundaries (agent and environmental boundaries) can be very difficult (Preiser at al., (2018).

In contrast, the theory of *autopoiesis*, the properties of living things, provides a different perspective. In this thinking, entities and agents are fully *closed* systems, and interactions with other agents and their environment are 'perturbations' that trigger changes within the bounded entity (whether an agent or a system). Recurring interactions eventually lead to 'structural coupling', which can be at a linguistic level for higher-order animals (Maturana & Varela, 1987). The biological explanations are included here as they are widely accepted explanations of agent adaptation within 'natural' settings, including social animals and humans. Conformance with natural CAS principles is important to formally evaluate whether socio-technical organisations are genuinely CAS. It is not a simple matter to declare in favour of CAS being open or closed systems. Where are the boundaries? Does our conception of ourselves as agents include our microbiomes? The question of open or closed system boundaries must always have the level of granularity identified (Mitchell, 2009; Preiser et al., 2018).

Management theories tend to follow prevailing scientific theories (Dooley, 1997). Mechanistic theories of organisational management such as Taylor's *theory of scientific management* dominated organisational thinking for much of the twentieth century. Efforts to find alternative theories naturally arose in response to the suboptimal and at times unpredictable reality of organisational performance. Organisational theorists explored the organismic system ideas of Bertalanffy, cybernetics, population ecology — variation, replication, and selection, and ultimately survival of the best fit with the environment. A commonality shared by these ecological theories is the idea of holism or irreducibility. Organisational behaviour should be viewed as a whole, rather than as component parts. Advances in the understanding of brain function as information processing also developed throughout the twentieth century. The notion of internal mental models, or

schema, to interpret information and action has implications for how groups of agents communicate within an organisation. Dooley (1997) emphasises the orders of learning proposed by Argyris and Schon. *First-order* learning is a comparison of perception with expectation via the internal schema. *Second-order* learning requires altering and updating the internal schema with new knowledge (Argyris, 1976; Dooley, 1997). Learning adds a mechanism for adaptation of multiple agents via information processing and updating schema in response to environmental signals. This conception of CAS is very close to Holland's pithy description of signals and boundaries as the building blocks of CAS (Dooley, 1997; Holland, 2012).

Assessing the scope and effectiveness of the use of complexity principles within social science is naturally of interest to researchers. Turner and Baker (2019) survey the origins of complexity science including CAS and add definitions and descriptions directly from literature sources. CAS properties are drawn together into their *tenets of complex adaptive systems*: (i) *path dependent* (ii) *systems have history*; (iii) *non-linearity;* (iv) *emergence;* (v) *irreducible;* (vi) *adaptive;* (vii); *operates between order and chaos*, and (viii) *self-organising.* From within this list, the notions of *path dependence, system history* and the idea of an *operational state* have not yet been discussed.

These three concepts are related. Path dependence is the idea that options for change are not infinite but bound by the available possibilities contained within the starting point. Previous option taking determines future options. System history is part of dynamical systems behaviour — systems that change can be said to move through different system 'states'. Much like the aphorism *never stepping in the same river twice*, a complex system is not static and today's system state may not be same as previous, or future states. Finally, the idea of order/disorder and chaos has also come from dynamical systems theory. For systems that are not static, the progression through different states can vary from stable within a narrow range (e.g. homeostasis) to complex (non-linear patterns) and through to chaotic (fully random).

Two final ideas complete the exploration of socio-technical CAS. First is determining the level of adaptation. It is important to consider whether adaptation occurs at the agent-level, or at the system-level. There is an increasing continuum of sophistication and time-scale through *learning, adaptation* (applying new knowledge) *and evolution* (persisting in a new state). These changes occur at the agent level, and system-level adaptation only

occurs when groups of agents change in a correlated manner (Carmichael & Hadžikadić, 2019; Wilson & Kirman, 2016). Second, the role of hierarchical leadership, ubiquitous within our social systems, needs to be explained. Incorporating leadership into CAS may stem from our natural human tendency to focus on hierarchical structures even in complex systems, for example queen bees in a bee colony, alpha leaders in a social mammal pack, or civic leaders in growing city (Carmichael & Hadžikadić, 2019). Taking time to think about the massively parallel information processing required quickly reveals that leadership or social dominance cannot equate to controlling all agent activity.

A relatively stable set of CAS characteristics as described within the literature can now be summarised. Table 3.3 presents a 'superset' of elements, observable features, or behavioural functions that distinguish CAS within the broader domain of complexity thinking.

Table 3.3: CAS Elements, Functions and Characteristics

Feature	**Elaboration**	**References**
Networks of connected agents	CAS consist of groups of agents, relationally connected in some form.	(Buckley, 2017; Holland 2006, 2012; Preiser et al., 2018).
Numerosity	CAS have high numbers of agents, interactions, and changes.	(Holland, 2006, 20012; Kauffman, 2000; Mitchell, 2009; Wiesner & Ladyman, 2019).
Not at equilibrium	CAS are Dynamical systems, operating on a continuum between maintenance of a stable state (or range of stable states), through complex and chaotic states. CAS function between chaos and stable ('the edge of chaos').	(Mitchell, 2009; Turner & Baker, 2019, p. 9).
Feedback	System outputs and interactions affect subsequent system inputs.	(Buckley, 2017; Carmichael and Hadžikadić, 2019).
Non-linearity	System outputs may not be proportional to inputs or may be discontinuous (step changes).	(Buckley, 2017; Wiesner & Ladyman, 2019).

Feature	Elaboration	References
Emergent higher order effects	System outcomes (e.g. an 'improvement') manifest as properties of the whole system, created by the interactions of system agents.	(Holland, 2012; Kauffman, 2000; Mitchell, 2009).
Agent schema	The internal mental models that agents use for perceiving the environment and determining actions.	(Carmichael and Hadžikadić, 2019; Dooley, 1997; Holland, 1992, 2006, 2012).
The ability to learn	Receiving new knowledge.	(Dooley, 1997).
The ability to adapt	Taking action based on new knowledge.	(Argyris, 1976; Dooley, 1997).
Coevolution	Agent interconnectedness means evolution (successful adaptation) may not be possible unless related agents also evolve.	(Holland, 2012; Preiser et al., 2018).
Self-organisation	Correlation of agent behaviour based on internal rules rather than centralised control.	(Carmichael & Hadžikadić, 2019; Holland 2006, 2012).
Irreducible	A CAS cannot be meaningfully broken down into constituent parts.	(Gershenson, 2013; Turner & Baker, 2019, p. 9).
Temporality	A CAS is dynamic and constantly changing over time.	(Turner & Baker, 2019, p. 9).
System history	Changes over time mean that CAS have history and path dependence (Options for change depend on the system history).	(Buckley, 2017; Turner & Baker, 2019, p. 9).

Although the emphasis varies, the literature on applying CAS thinking to other domains seems to be very consistent in identifying CAS elements and functions. In concluding this section, sufficient information has now been compiled to answer some elementary questions frequently asked of complex systems, for example, what is a

system? When does a system become a *complex* system? What makes a complex system a complex *adaptive* system? Systems are composed of interconnected parts, the structure and especially the interaction of those parts are important, and the idea of holism applies — the system is somehow more than the sum of its parts and not easily deconstructed. Structure and relationships can become complex through sheer number and type of elements, lack of equilibrium and feedback loops. A complex *adaptive* system can self-organise to make successful changes in response to changes in its environment. On that understanding, a large socio-technical system such as a hospital can easily be a system, a complex system, or a complex adaptive system, all at the same time, depending on where and when we observe it.

Evolutionary and ecological concepts feature strongly in CAS — *variation, selection, adaptation*, *cooperation,* and *coevolution*. The evolutionary/ecosystem foundation distinguishes CAS from the purely mathematics and physics-based complexity notions (e.g. Chaos theory). It is also clear that applying CAS ideas to human activity is an integral part of the CAS tradition, stretching back to general systems theory and continuing to the present day. Critically examining the justification for that enthusiasm is a potential research gap. A limited view of agency, together with ambiguity as to what self-organisation looks like in a hierarchical structure, pose the most serious challenges to acceptance of CAS ideas in socio-technical contexts. We may accept that a rainforest does not have 'leaders' or an overarching 'purpose', but it is a different matter to suggest that our socio-technical systems do not have hierarchy or some level of intentional design.

3.4. Synthesis 3: Quality Management and Complexity

The perspective in the healthcare QI literature that traditional quality management methods are not equipped to deal with complexity deserves more scrutiny. This section shows that responding to complexity has long been part of quality management and general systems thinking. Furthermore, methods such as Lean, perceived as reductionist by some healthcare QI proponents, continue to develop and incorporate ideas from complexity research. The knowledge domains of management and organisational theory also address improvement and quality improvement in the general sense. Several complexity-informed change frameworks being applied in healthcare systems originated within these non-healthcare fields. The key theme of *managing uncertainty* is apparent within all the approaches. Feedback paths, sensemaking, learning, holism, and cooperation all feature with varying emphasis as responses to the presence of uncertainty.

Cybernetics, the science of 'communication and control', has already been introduced. 'Organisational cybernetics' extends the basic principles beyond engineering and into the socio-technical realm (Beckford, 2002). The *Viable System Model* is a cybernetic system of modular, recursive feedback structures necessary to ensure any organisation can remain 'viable' as its environment changes. Crucial to the viable system model is the concept of maintaining requisite variety. Changes in operational context are inevitable and these changes need to be efficiently signalled to the appropriate operational, coordination or governance layer to respond appropriately. (Beer, 1984). The model is open to criticism, with a reliance on feedback signals being generated and distributed when and where they are required, and a failure to consider the role of organisational power dynamics (Beckford, 2002).

The subjective nature of system observation and interpretation is a barrier that any systems thinking framework must overcome. Diverse interpretations are recognised and elevated by *Soft Systems Methodology* (SSM). A 'hard' system stance assumes a system whose constituent parts can be understood and engineered. A 'soft' system stance acknowledges complexity and uncertainty. Therefore, a learning process must be established prior to intervention (Checkland, 2000). SSM seems to fit well with social knowledge domains such as healthcare, where multiple and conflicting points of view need to be reconciled. Note that the approach is not relativist in an epistemic sense; it simply acknowledges multiple subjective viewpoints and the need to discover what is really happening. SSM continues to be influential in healthcare improvement thinking (Augustsson, Churruca, & Braithwaite, 2020).

Closely related to the learning process integral to SSM is the concept of *Sensemaking.* As the name implies, if there is a gap between expectation and perception, we want to try and make sense of (understand) that gap. As with QI, sensemaking can be used in a very general sense or used to cover formal methods to reduce uncertainty. Sensemaking means taking deliberate action to understand, in a systematic and socially connected manner (Weick, Sutcliffe & Obstfeld, 2005). Perhaps the most mature sensemaking framework for improvement is the *Cynefin* framework introduced by Kurtz and Snowden (2003). Cynefin represents systems as moving dynamically between 'ordered' and 'unordered' domains, further broken down into *clear*, *complicated, complex* and *chaotic* system states. In this framing, 'order' means sufficient constraints to maintain system function, and disorder is an absence of organising constraint. Cynefin proposes different leadership

actions, decision-making and intervention actions based on the current system state and the level of knowledge. The essential insight is that the optimal management responses differ across the different system states. Cynefin has gained popularity internationally and is perhaps on the way to becoming a fully codified management method, with its own experts and consultants (Jackson, 2020). A potentially underdeveloped facet of Cynefin is the emphasis on 'ontological' complexity (what is actually happening) over 'cognitive' complexity (how difficult it is to try and establish what is happening and to reach a shared interpretation) (Jackson, 2020).

The common thread connecting the literature discussion thus far is the critical role of *learning*. An ability to learn is foundational to achieving quality outcomes in healthcare or any other knowledge domain. The generic notion of an *organisation* that learns is present in the work of Argyris (1976), Beer (1984) and Checkland (2000), with roots going further back to Deming's system of profound knowledge (Beckford, 2000; Deming 2000; Flood, 1998; Senge, 2006). The *Learning Organisation* framework introduced by Senge (2006) draws much of this source material together into an integrated holistic model. Five 'disciplines' work together to achieve long-term sustainability (i.e. the ability to adapt to constant change). The five disciplines of Senge's learning organisation are: (i) system thinking; (ii) personal mastery; (iii) mental models; (iv) shared vision, and (v) team learning (Senge, 2006). The language of 'discipline' is used purposefully to imply that practice is required to learn through action, individually and collectively. The five disciplines form a 'whole' in the sense that all five are necessary, with systems thinking being the gateway to a new way of perceiving an organisation holistically (Senge, 2006). Senge's use of systems language is somewhat bounded, based on system dynamics and common emergent behaviour archetypes such as the 'tragedy of the commons' or 'shifting the burden' (Flood, 1998). These unwanted system outcomes are the products of dynamic interaction, with balancing and reinforcing feedback loops occurring amongst different system elements. Maintaining awareness of these system interaction archetypes is certainly important. However, additional complexity measures relevant to socio-technical systems, such as cognitive complexity, agent network structure, power dynamics between agents, and previous history, are underexplored.

The 'mental models' as described in Senge's framework relate to persistent or recurring thought patterns in use (Flood, 1998). A less prescriptive use is introduced below, which simply describes agents' internal expectations and perceptions of a given

context. These distinctions aside, the learning organisation has genuine potential as an integrated, complexity-informed improvement framework. Intriguingly, applied use of learning organisation principles within healthcare quality improvement appears to lag behind the other approaches described in this section. Applications have also tended to be at the organisational/facility level, rather than QI project or team level (Akhnif, 2017). Potentially, the substantial re-orientation required of an organisation is more challenging in healthcare than for-profit organisations.

Redefining the role of leadership is a feature of contemporary management thinking. Complexity leadership frameworks challenge the inflexibility of conventional leadership paradigms (Heifetz et al., 2009; Uhl-Bien & Marion, 2009). Two contributions are apparent. 'Technical' activity (simple, known, or able to be known) is distinguished from 'adaptive' activity, where the required action is uncertain. 'Adaptive leadership' is the necessary *enabling* of discovery processes, iteration, and learning in a bottom-up manner from workers (Heifetz et al., 2009). This distinction is thus very similar to the clear/complicated/complex categories of Cynefin and the complexity leadership model (Uhl-Bien & Marion 2009).

The concept of teamwork is also prominent within organisational theory. The relevance to this research is the extent to which connected networks of agents within a CAS function as teams. How do agents work together towards their shared (or not shared) objectives? *Shared* mental models allow the transition from individual agents evaluating their internal mental schema against environmental signals, to groups of agents ('teams'). Team mental models are "*organised mental representations of the key elements within a team's relevant environment, that are shared across team members"* (Mohammed, Ferzandi & Hamilton, 2010, p. 876). Technical questions of definition, mechanism and measurement are acknowledged in the academic discourse on shared/distributed cognition and mental models. A pragmatic approach for this study is to begin with a focus on the gaps in shared representation, and working to close these (Cronin &Weingart, 2007; Wang, Mannix, & Cronin, 2016). In effect, this is another description of the sensemaking process.

Production-oriented improvement methods areas such as Lean production and supply chain logistics have their own growing literature on complexity. The limitations of prescriptive instructions, and the danger of isolating individual system components are

acknowledged. Non-linearity in complex systems can hide downstream consequences, and just-in-time systems can be paralysed if insufficient slack is built into the process. Balancing these vulnerabilities, pull production, visual management, and kaizen have all been shown to increase resilience under uncertainty (Ferreira & Saurin, 2019; Saurin, Rooke & Koskela, 2017; Soliman & Saurin, 2017).

As the plurality of systems methodologies has evolved, there is a temptation to view the different approaches to quality improvement as a well-stocked tool kit. Selecting the right approach and method for the specific condition is what matters. Jackson's (2001) *Critical Systems Thinking* proposes 'a system of systems methodologies' pragmatically acknowledging (i) that there are many types of system, and (ii) that all approaches have their strengths and weaknesses. There is no single correct approach, just the right approach for the current problem, and creativity and a holistic systemic perspective are necessary (Beckford, 2002; Jackson, 2001, 2006).

The preceding paragraphs demonstrate that the quality management domain, together with the adjacent domains of organisational and management theory, are all well acquainted with complexity. There is a history that continues to evolve and develop in response to the challenge that complexity presents to organisations. As noted in Table 3.2, methods from these fields are making contributions to healthcare QI. However, the different prescriptions and solutions often retain an organisational lens including notions of hierarchy. The solutions tend to modify existing socio-technical approaches in response to complexity conditions. Improvement/correction/change is treated as a separate or external intervention on the system. There does not yet seem to be a quality improvement method built fully and exclusively on CAS principles.

3.5. Synthesis 4: Modelling Healthcare Quality Improvement

Common strategies can be observed across several varied examples of healthcare QI modelling. These models are a way to explore ideas not easily observed or tested empirically. The focus also tends to be on aspects of theory, at a high level of abstraction, as opposed to creating highly specified digital emulations of the real world.

Given its long history in complexity and systems thinking, *systems dynamics* is an obvious starting point for modelling complex change. This approach uses stock and flow diagrams to represent resources and activities moving and accumulating within a system. Causal loop diagrams show balancing (negative) and reinforcing (positive) feedback

paths, collectively representing the complex interaction of components within a system. Systems dynamics has been used in healthcare QI to examine the enablers and barriers within acute care, and to demonstrate the complexity of policy interventions in response to the COVID-19 pandemic (Muttalib et al., 2021; Sahin et al., 2020). Whilst undoubtedly useful for 'complicated' systems (i.e. systems theoretically able to be discerned and understood by experts), systems dynamics seems to rely on a level of structural fidelity that may not be possible to discern or verify for many complex systems.

The tools of network science offer access into the roles of agents and agent interactions within CAS. The impact of specific agent network features such as centrality, node density and homophily can be explored using network analysis measures (these measures are defined further in section 6.6). Network science concepts are valuable in building up an understanding of the barriers and enablers to communication and coordination amongst different healthcare teams. (Carney et al., 2015; Levula et al., 2013; Manukyan, Eppstein, & Horbar, 2013).

The dominant modelling approach used to study CAS is Agent-based modelling (ABM). This approach is especially suited to complex systems involving autonomous agents (Holland 2006; Preiser, et al., 2018). The focus on individual agents allows the system complexity to be examined from the granular level of the agent activity, rather than attempting to faithfully define the full scope and detail of high complexity systems 'from the outside in'. ABM usually incorporates stochastic procedures to reflect levels of uncertainty and randomness in the system (Holland, 2012). In other words, ABM are not deterministic models, although repeatable patterns may emerge for a given set of input conditions.

Studies specifically applying ABM to quality improvement are still rare. Two recent examples explore the effectiveness of public health interventions under complexity and how knowledge levels increase within a clinical network. Ornstein et al. (2020) applied the theory of 'rugged landscapes' from evolutionary biology to show how public health implementation efforts might be sub-optimal when not matched to the level or type of complexity present. From the *learning health systems* field, Seid et al. (2021b) modelled how patients and clinician interactions influence engagement levels and recursively increase the overall level of knowledge available within the system. Both examples are exploratory and at an early stage of maturity. Nevertheless, they demonstrate a viable

research path for aspects of healthcare delivery that are otherwise difficult to access. Recent guidance for ABM simulation modelling in public health research demonstrates the growing acceptance of ABM in healthcare research. As with any computational modelling, the need for good discipline with model specification, adequate access to data, and good validation practices are emphasised (Badham et al., 2018). The development of the ABM for this research is detailed in Chapter 6.

3.6. Scoping Review

3.6.1. Approach

The goal of the scoping review was to identify a comprehensive, relevant pool of secondary data to be used for empirical assessment of the conceptual model developed in this research. A scoping review was selected as the appropriate category of comprehensive literature review. Indications for when a scoping review may be the most appropriate choice include: (i) uncertainty over the categories and content of the existing literature; (ii) the need to clarify key concepts and characteristics, and (iii) identifying and analysing knowledge gaps (Munn, et al., 2018). Each of these indications apply to the present study. Scoping reviews still retain structural features to ensure comprehensiveness, reproducibility, and minimise bias, and the PRISMA standards contain an extension for scoping reviews (Tricco et al., 2018). There is also precedent for scoping reviews in the specific context of healthcare improvement research. Two recent and relevant examples include: (i) a scoping review protocol for the literature on soft systems methodology within healthcare (Augustsson, Churruca, & Braithwaite, 2018), and (ii) a scoping review on the formation of quality circles within primary care (Rohrbasser, et al., 2018).

3.6.2. Review Structure

The framework developed by Arkey and O'Malley (2005) was used to structure the review. The five framework stages are (i) Identifying the research question; (ii) Identifying relevant studies; (iii) Selecting studies; (iv) Charting (presenting) the data; and (v) Collating, summarising and reporting the results (Arkey & O'Malley, 2005).

3.6.3. Identifying the Research Question

This study examines a specific research question: *What system factors influence the success or failure of healthcare QI activity within complex adaptive systems?* To formally review and map the literature relating to this question, the scoping review covered the

published literature relating to the *intersection* of healthcare, quality improvement and complex adaptive systems. The intersection is intentional, given the specificity of the topic: quality improvement, as practiced within a healthcare setting, where the context is recognised or theorised as a complex adaptive system. The scoping review domain intersection is represented in Figure 3.1:

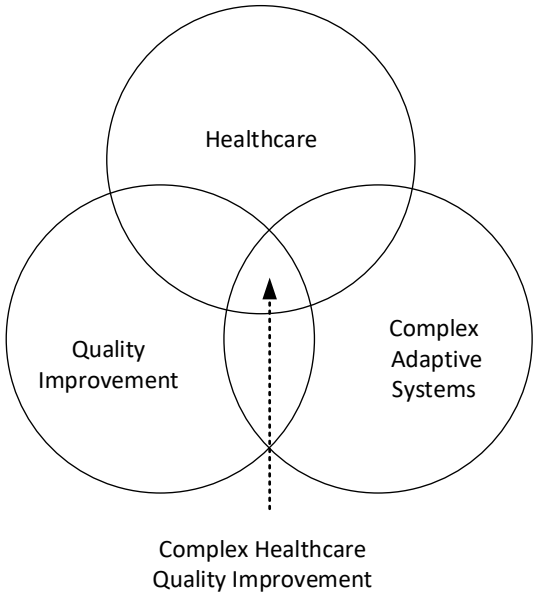


Figure 3.1: The Intersecting Literature Domains of the Scoping Review

A sense of the scale of the contributing domains is provided by Table 3.4, showing the indicative size of the literature from one major database (Scopus):

Table 3.4: Literature Quantity for the Contributing Domains

Keywords	**Scopus search results (09/01/2020)**
"Quality improvement"	1,142,660
"Healthcare"	1,428,475
"Complex adaptive systems"	521,523

Focusing on the intersection between these domains thus ensures the review task is achievable; much of the extensive literature in each domain is not relevant to the research question.

3.6.4. Eligibility Criteria

The review was limited to English language scholarly literature, without limitations on the category of literature, to be as comprehensive as possible. Peer-reviewed academic journal articles, books, and grey literature (theses, conference presentations and proceedings) were all in scope.

3.6.5. Identifying Relevant Studies

Seven electronic databases were searched. This search strategy resulted in a very high number of initial duplicates, but the range of databases was deemed important to cover the interdisciplinary range of the topic. Searching key databases from different academic fields ensured adequate coverage across the literature in healthcare, operations management, education, public policy, engineering and science:

1. CINAHL Complete
2. PsycINFO
3. PubMed
4. Web of Science
5. Business Source Complete
6. Scopus
7. IEEE Xplore

Preliminary search strategies were trialled using the Massey University 'Discover' database, which offers a google-style search of the university's main databases including the above sources. The initial exploratory searching led to the selection of the keywords in Table 3.5 that follows. For example, 'complexity' on its own is too generic even when restricted to healthcare. 'Complexity theory' and 'complexity science' are used interchangeably in the literature but searching on only one returns incomplete results. 'Complex adaptive systems' also returns a different set of results and does not identify relevant work described only as 'complex systems'. 'Quality improvement' is the established term within healthcare improvement literature (as opposed to 'continuous improvement' or just 'quality' or 'improvement'). Because the research is examining quality improvement methods, the four most common methods as described in the healthcare QI literature are also searched in their own right.

Table 3.6 lists the Boolean phrases used to search title, abstract and keywords of items within the identified databases. A two-stage approach to the search was adopted to ensure the search queries were not overly complicated and hence easily reproducible. The first level of query identified relevant literature relating to QI and complex systems. A second search then identified the subset of the level one search results that relate to 'health' OR 'healthcare'. This approach also allowed for relevant review of the adjacent literature of QI and complex systems but outside the healthcare domain.

Table 3.5: Keyword Search Query

Keywords (searched in Title, Abstract and Keywords)
Level one: database search
('quality improvement' AND ('complex systems' OR 'complexity theory' OR 'complexity science'))
(('PDSA' OR 'plan do study act') AND ('complex systems' OR 'complexity theory' OR 'complexity science'))
(('lean' OR 'six sigma') AND ('complex systems' OR 'complexity theory' OR 'complexity science'))
('improvement collaborative' AND ('complex systems' OR 'complexity theory' OR 'complexity science'))
Level two: search of valid level one results
('health' OR 'healthcare')

The Boolean search phrases were tested for each database, including case sensitivity. The specific syntax of each search is provided in Table 3.6:

Table 3.6: Search Query Syntax per Database

Database search syntax (searched in Title, Abstract and Keywords)	
CINAHL Complete	“quality improvement” AND (“complexity science” OR “complexity theory” OR “complex systems”)
PsycINFO	(‘quality improvement’ AND (‘complex systems’ OR ‘complexity theory’ OR ‘complexity science’))
PubMed	"quality improvement"[All Fields] AND ("complexity science"[All Fields] OR "complexity theory"[All Fields] OR "complex systems"[All Fields])
Web of Science	TS=("quality improvement "AND ("complexity science" OR "complexity theory" OR "complex systems")) Databases= WOS, BIOABS, CABI, CCC, FSTA, KJD, MEDLINE, RSCI, SCIELO Timespan=All years Search language=Auto
Business Source Complete	quality improvement AND complex systems OR quality improvement AND complexity theory OR quality improvement AND complexity science
Scopus	TITLE-ABS-KEY ("quality improvement" AND ("complexity science" OR "complexity theory" OR "complex systems"))
IEEE Xplore	"All Metadata":"quality improvement" AND ("complex systems" OR "complexity theory" OR "complexity science"))

3.6.6. Selecting Studies for Review

The PRISMA guidelines flowchart (Moher et al., 2009) has been used to represent and record the identification, screening and selection process (Figure 3.2). The number of databases searched resulted in a high number of duplicates, of which many could be eliminated by automated matching using MS Excel. 633 articles were excluded at screening for one of the following reasons: (i) highly detailed scientific, technical or clinical content without reflections/analysis on complex systems or improvement methods in general; (ii) QI reports/cases or reports that lacked wider context on complexity or improvement thinking; (iii) short or informal opinion articles not contributing new content or analysis; (iv) spurious search results, and (v) manually identified duplications. 183 articles were identified for full text eligibility assessment, of which 93 were excluded with reason. Of these 93, 34 articles were reviews or theoretical perspectives relating to healthcare QI and complexity, as opposed to empirical QI cases. Of course, these theoretical perspectives are still highly relevant and have been incorporated into the overall literature synthesis. 59 articles were also excluded at full text reading due to failing to meet the inclusion criteria. This left a balance of 90 articles for inclusion in the qualitative analysis.

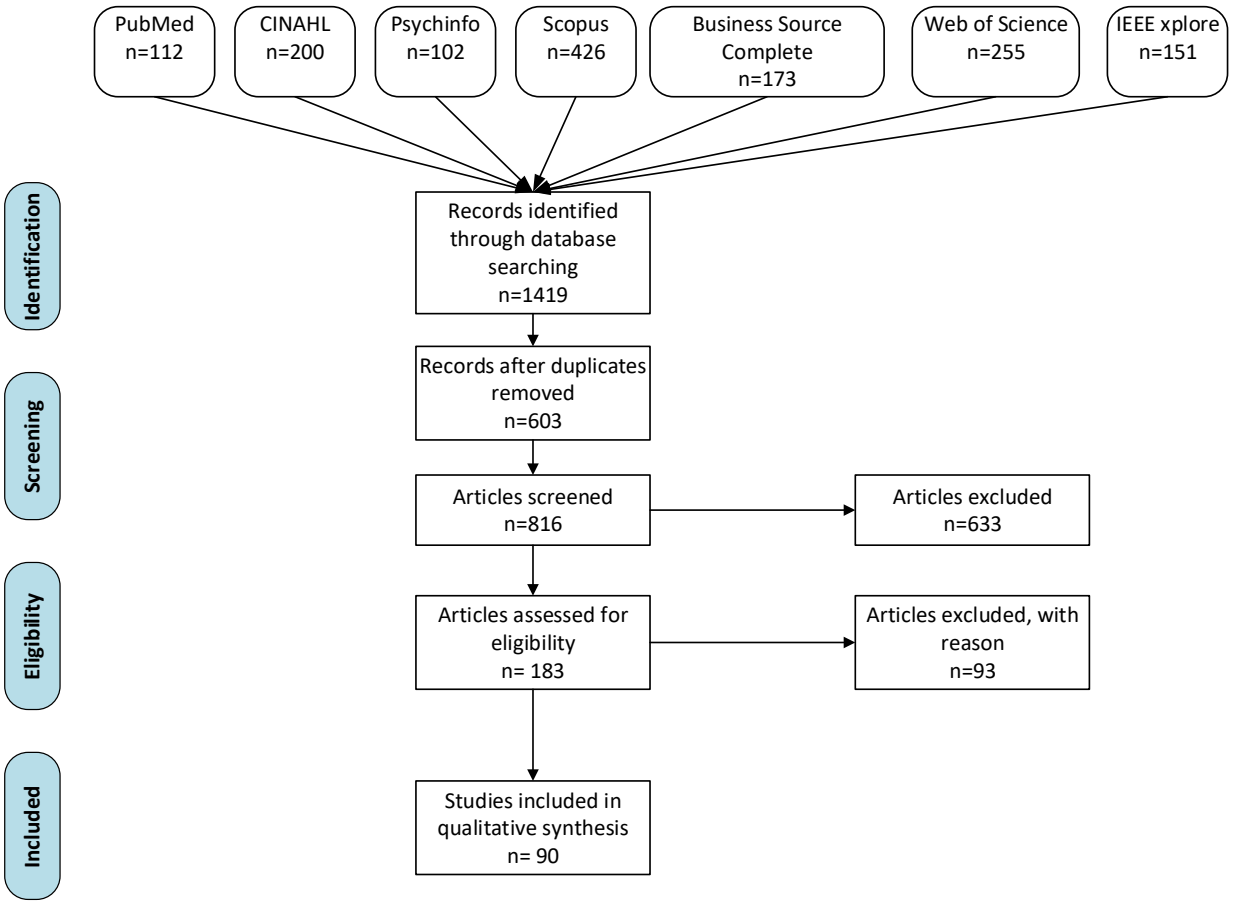


Figure 3.2: Scoping Review Selection Process (Adapted from Moher et al., 2009)

3.6.7. Scoping Review Effectiveness Assessment

The scoping review results provide confidence in the search strategy and selection. Firstly, the research 'zone' certainly appears to be legitimate, in that the targeted intersection of quality improvement, healthcare and complex adaptive systems is neither too restricted nor excessively general. The database keyword search strategy was effective, as was the decision to search multiple databases. These have ensured the net has been cast wide across inconsistent use of key terms (e.g. complexity science vs complex systems vs complex adaptive systems). The results also confirmed the breadth of complex systems work across disciplines — fundamental sciences, engineering, health, management, social policy, and education. Medicine dominates, but this reflects the choice of databases used.

Key authors and papers that were expected based on informal scans of the literature review or prior searching were picked up in the search. Where prominent authors on the topic were missing from the search results, the omission was explainable based on the search query, for example the focus of many authors on complex systems in health is at

the system transformation level, with no explicit reference to quality improvement and complex systems in the same article. Articles not meeting the eligibility criteria but potentially useful to the research were also identified. The relevance of this group is in assessing existing frameworks and tools to measure and understand complex adaptive systems. The work of Tarcisio Saurin and colleagues on positioning Lean in complex systems (Ferreira & Saurin, 2019; Saurin, Rooke, & Koskela, 2013), and Yaneer Bar-Yam (Bar-Yam, 2004), on general models of complexity are highly relevant, even though they fell outside of the eligibility criteria as strictly applied. These works have been incorporated into the relevant literature synthesis or the discussion as appropriate. Descriptive statistics and detailed qualitative analysis of the 90 empirical studies follows.

3.6.8. Descriptive Statistics

Table 3.7 shows the count of articles included in the scoping review, by year of publication. The list spans 21 years, with 50 of 90 articles from 2018 or later.

Table 3.7: Year of Publication and Complexity Terminology Usage

Year	Count of Articles
2022	4
2021	6
2020	12
2019	14
2018	14
2017	8
2016	6
2015	4
2014	3
2013	5
2012	2
2011	3
2010	2
2009	3
2007	1
2004	2
2001	1
Total	**90**

Complexity Terminology*	Count of Usage**
High number of elements	36
High number of interactions	32
Diversity of elements	30
No Ref/Colloquial Ref only***	27
Complexity science	24
Self -organisation	22
Complex adaptive systems	19
Non-linearity	15
Emergence	14
Non-equilibrium	13
Feedback	11
Adaptive behaviour	9
Memory	7
Attractors	4
Total	**263**

***** As explicitly described or apparent from description of context

****** Multiple references per case; total greater than number of cases

******* Use of ‘complex’ or ‘complex system’ with minimal further elaboration

Table 3.7 shows the usage of complexity terminology across the 90 articles. Wiesner and Ladyman's (2019) conditions and products of complexity were used for categorising the diverse and frequently ambiguous references to 'complexity'. Slightly more than one third (98/263) of the complexity references were to large numbers of people, system components, interactions, or high diversity/heterogeneity of these factors. As reported by these practitioners, complexity is strongly associated with the quantity and variation of system elements.

There were 27 references which did not meaningfully expand on 'complex', 'complex systems', or 'complexity' once these terms had been introduced. These authors perhaps assumed their usage required no further explanation, or that general assertions such as 'healthcare is complex' would be accepted at face value by readers. References to *complexity science*, *self-organisation* and *complex adaptive* systems made up 25% (65/263) of the complexity descriptions. *Non-linearity*, *emergence* and *non-equilibrium* (lack of stability; high rates of change) were the most common technical terms used, with 16% (42/263) of the usage. There was limited appearance of other complexity measures, with just 12% (31/263) for *feedback*, *adaptive behaviour*, *memory* or *attractors*.

The range of improvement methods and approaches in use across the 90 cases was revealing. Of course, the results reflect the original search criteria, which explicitly searched for *PDSA, Lean, Lean Six Sigma* and *Healthcare Collaboratives*. Nevertheless, Figure 3.3 justifies the frequency assumptions from the literature review. Once *Implementation Science* and *Resilient Healthcare* are included, the dominant frameworks in use are clear. Encouragingly, there were also 5 cases where complexity thinking, including CAS, was framed as the improvement *approach* (as opposed to the context or an analytical lens). The red line on Figure 3.3 tracks a Pareto summary (seen on the secondary vertical axis), with the 80% mark reached at 'Network science'. Figure 3.3 then displays a long tail of improvement approaches, methods and tools, indicating the willingness of researchers to explore new improvement thinking, whilst also suggesting that none of these alternatives have yet become widespread.

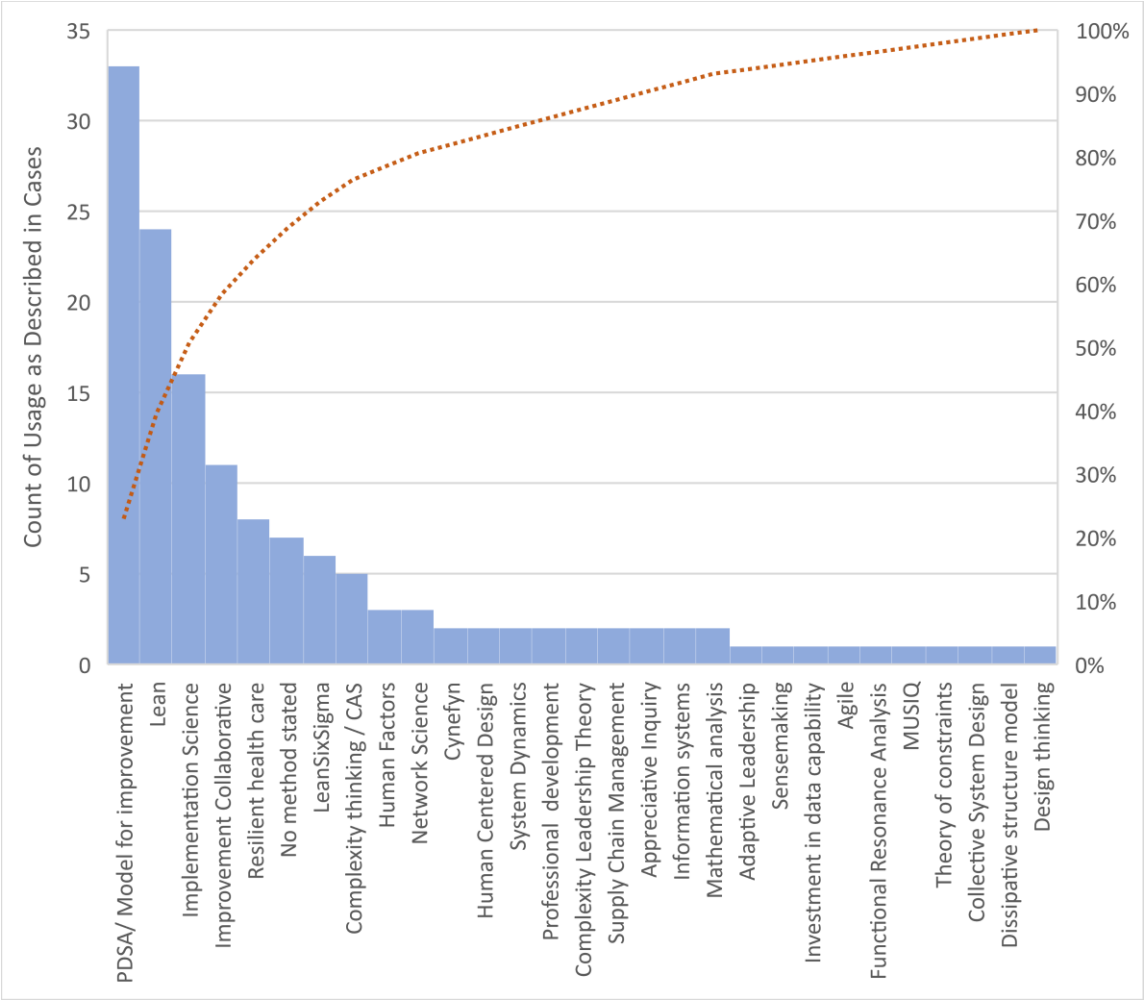


Figure 3.3: Improvement Methods Identified in the Secondary Data Case Studies

Appendix 2 contains further summaries of the secondary data by country, healthcare focus area, and journal. The cases spanned 41 healthcare focus areas across 22 countries, with the United States, United Kingdom, Canada, and Australia accounting for 64 of 90 cases. There was a very wide dispersion of journal (65 different journals). After examining the initial range of content and confirming relevance to the research question, the selected articles were ready for in-depth examination. As part of the research design, the secondary data case studies were used to assess empirical validity of the conceptual model and simulation findings, rather than contribute a priori knowledge to develop the model (although additional findings are incorporated). For this reason, the analysis is presented in Chapter 8, following the simulation results.

3.7. Literature Review Summary

General change strategies within healthcare have been outlined, with each of these strategies representing a different aspect or focus of healthcare quality improvement (used

in the general sense). Evidence-based medicine, learning health systems, clinical safety, resilience engineering, and service evaluation all coexist to some extent with the narrower concept of QI methods — defined as those improvement approaches adopted and evolved from traditional quality management. It was noted that all the change strategies must contend with complexity, and that they are not mutually exclusive. The building block of QI, the PDSA cycle, is often used as a rapid learning process within other healthcare improvement methods.

Questions over the effectiveness of healthcare QI have been raised in the literature. The dominant argument, beyond questions of cultural fit, is that QI methods are seen to represent reductionist and deterministic approaches not suitable for the complexities of healthcare. A considerable body of literature posits that healthcare is a CAS, and effective management, including improvement, needs to be oriented to CAS principles. According to this argument, change should be viewed as an emergent property of localised agent activity, and not imposed in a top-down manner. This is claimed to be the only way of meaningfully managing the high complexity conditions that exist — high numbers of agents and interactions, multiple feedback loops and non-linear outcomes all giving rise to high uncertainty. A set of identifying features of CAS from across the literature were set out in Table 3.3 Distinguishing CAS from 'complexity' more generally, *connected groups of agents*, *learning*, and *adaptation* are the key attributes present in CAS not shared with other forms of complex systems.

To offset any perception that conventional quality management methods have not considered complex system thinking, notable strategies within this knowledge domain were outlined. Important ideas from management/organisation theory were also included, particularly *sensemaking* and complexity leadership models. *Soft Systems Methodology*, the *Cynefin* sensemaking framework, and *Complexity leadership*, are three examples of frameworks from the management/organisational fields that have appeared in the healthcare complexity management literature. Because this research is using simulation modelling to build and test theory, recent approaches to modelling QI were surveyed. The specific intersection between simulation modelling and QI is still very small but the number of examples is growing.

Finally, a research gap in the literature is now proposed, noting that the wide advocacy for healthcare as a CAS has not come with a detailed explanation of how

sophisticated human socio-technical systems function in formal accordance with CAS principles.

3.8. Research Gap

In socio-technical settings, the challenge of constructing an integrated framework for change is complicated by *human agency* and higher-order cognition. Paley and Eva (2011) contend that cognition in human agents, and the role of cognition in self-organisation, results in a completely different process than for system agents who are not self-aware. The CAS principle of agents following simple rules without an overall design is seemingly in conflict with the rich and varied decision making we carry out every day. Human awareness also allows agents to be both participant and observer in a 'system', a feedback mechanism not present in natural CAS (Goldspink & Kay, 2003). Paley and Eva's concerns do not seem to have been explored further in the literature, although Khan et al., (2018) suggest they are still part of the 'healthcare as CAS' knowledge domain to be filled in. If healthcare is indeed a CAS, then it should be possible to develop specifically CAS explanations of healthcare activity, including healthcare QI.

A tension exists between maintaining a relatively stable set of operational states within a system, versus the system generating sufficient variety to successfully adapt (Braithwaite, 2018; Junghans, 2017). Together with the issue of agency, this tension seems to be at the heart of the 'centralised/top-down control is bad' versus 'local adaptation/self-organisation is good' language within the literature. At one level the argument is about having sufficient autonomy to modify ideas locally given the variety of local contexts. But at another level, there is a challenge to the notion of a hierarchical structure built for stability, now needing to survive in the presence of high complexity. Any effective explanation of socio-technical CAS will need to show how desired, macro-level system outcomes (a system *purpose*) is facilitated/inhibited by organisational structure but produced directly from independent agent actions.

The literature shows that strategies and frameworks for managing change within complexity exist, including some specifically built for the healthcare setting. These are undoubtedly helpful to QI practitioners, and many are in active use within healthcare (Anderson et al., 2015; Augustsson et al., 2020; Bäcklander, 2019; Kringos et al., 2015; Perla et al., 2013; Snowden & Roncati, 2021). However, strong socio-technical filters are present in all these frameworks. Leadership strategies, leadership actions, organisational

structural levels and prescriptions for appropriate actions are provided. A genuinely CAS explanation of healthcare QI should be able to explain the mechanism of system change from the granular level of agent activity where it occurs. Illuminating CAS change mechanisms in socio-technical settings may give momentum to fully embracing CAS principles and allow organisations to flourish in the presence of complexity. Chapter Four now introduces the conceptual model, developed from CAS and QI literature, which is the starting point for closer examination of the research gap.

Chapter 4 Conceptual Model

4.1. Introduction

This research develops, refines, and tests a theory of change, specifically a CAS explanation of QI. The conceptual model introduces the theory in a form that can be scientifically evaluated. The constituent elements are derived from CAS and QI literature; therefore, the conceptual model follows on from the literature review and forms the concluding chapter to Part One of the Thesis. In this chapter, the components of the conceptual model are introduced and explained. The component parts combine to form three functional levels within the model. Section 4.2 covers the first level of the model, networks of adaptive agents. Problem solving activity forms the basis of level 2, and this is the focus of section 4.3. The final level is system regulation, described in section 4.4. Following the outline of the model levels, a minimum specification for the conceptual model is proposed in section 4.5. The *complex quality improvement network* (CQIN) is then presented in section 4.6. The chapter concludes with an assessment of the CQIN conceptual model as a theory of change.

4.2. Level 1: Networks of Adaptive Agents

The initial level establishes the CAS within the socio-technical context. This study proposes that QI within CAS is a special case of CAS. There is a conflict between the intuitive appeal of interpreting complex social organisations as CAS and the valid question of system design within CAS. As a starting point, it is taken as axiomatic that at given moments in time, CAS agents involved in healthcare QI do coordinate their actions towards a goal, and that goal can be defined and understood at a level beyond one individual agent. In these circumstances, the interpretation of "CAS" is important. In this usage the CAS is not the clinical microsystem, project team, organisation or population group that first comes to mind. The CAS is a *specific collection of agents within these entities, cooperating on a specific goal*. This definition aligns with the two distinct CAS categories defined by Wilson and Kirman (2016). In CAS_1, *adaptation* takes place at the system level, for example, with the human immune system. CAS_2, in contrast, are interconnected groups of adaptive agents, who each respond adaptively to their environment and any system-level change emerges from this *agent-level* activity (Wilson & Kirman, 2016). Using this agent-level conceptualization of healthcare QI, the CAS in

focus becomes simply one of very many interconnected adaptive agent networks within a larger socio-technical system such as a hospital or healthcare facility.

Understood as a network of interacting agents, the interactions and their effects can be examined by appropriate network analysis methods. Network science, built on graph theory, provides tools that are especially useful to examine network structure. These tools allow the heterogeneity of the agents and the number and nature of the interconnections to be formally described (Barabási, 2016). The boundaries of any given agent network and the role of adjacent and interlinked networks also needs to be adequately accounted for in any explanation of a given network's behaviour. Some form of dynamic network boundary delineation is required as CAS are open systems without fixed boundaries (Holland 2006; Nair & Reed-Tsochas 2019; Umpleby, 2008). Where the CAS stops and the external environment begins is not fixed. Nair and Reed-Tsochas (2019, p. 86) describe this permeable boundary as the *enacted and interpreted boundary* within a CAS, dynamically linked to the external environment. This eloquent description incorporates both transience and subjectivity into the network boundary. Networks will change as agents join and leave. Implementing change may also require other related agent networks to adapt if change is to occur. In natural CAS, adjacent groups adapting in response to each other is understood as coevolution and this study retains the term (Holland, 2012; Kauffman, 2000). Describing inter-network adaptation as coevolution keeps the underlying evolutionary principles for change within CAS in focus.

4.3. Level 2: Problem Solving as Adaptation

CAS theory asserts that system-level change emerges from localised agent interactions where the agent behaviour conforms to a small number of internalised rules (Holland, 2012; Kauffman, 2000: Mitchell, 2009). Therefore, identifying the specific agent adaptation processes within a QI setting is necessary. A comprehensive articulation of CAS behaviour has been provided by John Holland (1992, 2006, 2012). Holland defines four essential CAS attributes: (i) *Parallelism*, where large numbers of agents interact by sending and receiving signals; (ii) *Conditional action*, where agent actions are conditional on signals and rules available to the agent; (iii) *Modularity*, where activities can be combined to form larger "subroutines", and (iv) *Adaptation and Evolution*, where the agents change over time, via adaptations that improve a situation (rather than random change). This description of adaptation is now examined more closely.

Holland (2012) proposes that adaptation requires *credit assignment*. This term refers to the weighting and rewarding of individual signals and rules within the parallel activity. Therefore, a means of detecting individual signals within parallel messaging activity is required, along with the ability to determine the relative importance of the signal. Having detected a signal and judged it requiring action, an agent then acts based on the conditional rules within its internal schema. When an agent's existing rules are incomplete or inadequate, a process for *rule discovery* is required. A computational biologist, Holland pioneered 'genetic algorithms' to perform this function. New rules are generated via random mutations to existing rule/solution possibilities and a fitness test applied. Successful rules are then retained in an updated schema via credit assignment. Holland's *learning classifier* system impressively encompasses signal detection, signal evaluation, rule evaluation, rule generation, rule selection and rule retention (Holland, 2006, 2012). Learning classifier systems and genetic algorithms have both been influential in solution optimisation and machine learning. (Urbanowicz & Moore, 2009).

Satisfactorily expanding or replacing this *adaptive* classifier 'kernel' presents a significant challenge to the proposition that natural CAS principles are appropriate for sophisticated human behaviour. Plausible, non-biological agent mechanisms are required to replace the evolutionary processes that occur across generations within natural (non-self-aware) CAS. What replaces the inherited and learned behavioural rules, genetic variation, genetic drift, and natural selection? (Holland, 2006, 2012; Kauffman, 2000). Conceptualising improvement as a *problem-solving process* provides a feasible bridge between the biologically encoded, rule-based agent behaviour within a CAS and the goal-oriented activity of human agents.

Adaptation in response to environmental signals implies closing a gap between one system state interpreted as requiring action, and a subsequent system state where the desired action has been taken. Selecting options to close this gap is exactly analogous to problem solving where solution options are taken to navigate a given problem space (the set of available operations that can be carried out). Explicit problem-solving strategies, such as *hypothesis testing*, *means-end analysis* or the *goals-assumptions-elements-operators* model provide a formal and non-biological mechanism to introduce *purpose*, *awareness*, and *agency* into the CAS adaptation process. Problem solving can provide the necessary learning and adaptation functions for a socio-technical CAS (Anderson, 2005; Cronin & Weingart, 2007; Doolittle 2014; Newell & Simon, 1972).

Cooperative problem solving establishes, tests, and evolves the internal mental models or *schema* of groups of participating agents (Cronin & Weingart, 2007; Dooley, 1997; Holland, 2012; Mohammed et al., 2010). Problem solving requires knowledge which, if not initially present in the agent network, must then be acquired (i.e. learned). Therefore, a learning gain from each learning activity is posited as necessary to acquiring sufficient knowledge within the agent network (Anderson, 2005; Deming, 1986; Doolittle, 2014; Reed & Card 2016). Learning leads to the creation of new rules, a necessary feature to translate Holland's classifier system to a goal-directed context.

4.4. Level 3: System Regulation

The problem-solving actions of agents are attempts at system regulation and control. To regulate a system, agents need to receive feedback signals. Signal paths are therefore necessary to receive and distribute the environmental signals. At one level, this is a straightforward application of cybernetic principles (Ashby, 1956; Umpleby, 2008). However, the role of agent schema and differing levels of agency need to be considered in more detail at this point.

More than just a list of conditional rules, agent schema can be viewed as the internal encoding or *map* of the environment that each agent maintains. Understanding and responding to environmental signals is thus a process of comparing the external environment with an internal model of that environment. This idea has been expressed formally within the 'Bayesian brain hypothesis', where the brain is viewed as an inference machine making predictions and testing these predictions with sensory data. The internal model is updated when new understanding is gained (Bain, 2016; Friston, 2010). Treating QI as an inherently Bayesian process is helpful for responding to complexity because complexity limits information and interpretability. Understanding a system is an iterative process of making predictions and updating beliefs as more information is obtained.

A valid objection to the cybernetic control description is that it implies a very low level of agency that is not suitable to describe human interactions, which are obviously more rich and nuanced than a simple stimulus and programmed response mechanism. However, this criticism misrepresents the sophisticated levels of control mechanism that Ashby (1956) proposed. Furthermore, evolutionary psychologists today recognise a hierarchy of agency levels allowing for working memory, reasoning, and social regulation (Tomasello, 2022).

4.5. Minimum Specification of the Conceptual Model

The minimum requirements for a model explaining QI within CAS are now specified. This conceptual model is parsimonious in that it is limited to the minimum set of constructs to model QI within CAS as problem-solving activity within networks of adaptive agents. A system capable of carrying out purposeful improvement actions requires: (i) the circulation of system information (signals) amongst a network of agents; (ii) some level of shared interpretation of issues, goals and assumptions related to an improvement; (iii) problem-solving capability and resources adequately matched to the problem and wider system context; and (iv) an ability to implement changes within the defined system as well as any closely related systems (agent networks) that may be impacted (Cronin & Weingart, 2007; Holland, 2012; Newell & Simon, 1972). Having set out the theoretical foundations for the conceptual model, the *Complex Quality Improvement Network* (CQIN) model of QI within CAS is now introduced.

4.6. The Complex Quality Improvement Network (CQIN)[3]

The CQIN conceptual model of a generic problem-solving network is presented in Figure 4.1 A network of agents (A) cooperates in deliberate, iterative problem-solving activity (Paths G-B), receiving and interpreting system feedback (paths B-E-F, and C-D-F). The agent interactions are moderated by the complexity of the setting (H) and exogenous factors from the external environment (I). A learning gain is achieved per iteration.

[3] The author is aware of one earlier usage of the CQIN acronym. As referenced in Adolf & Custer (2002), this earlier usage of the acronym refers to a *continuous quality improvement network.* The context is a collection of US Colleges using self-assessment tools for continuous improvement, as part of their institutional accreditation requirements. The term refers to continuous improvement tools, projects and maturity assessment within a Baldrige-criteria quality framework. There are no references to complexity, CAS or agent networks. Apart from the shared high-level domain of quality improvement, there is no relationship between this early usage of the CQIN term and the current research.

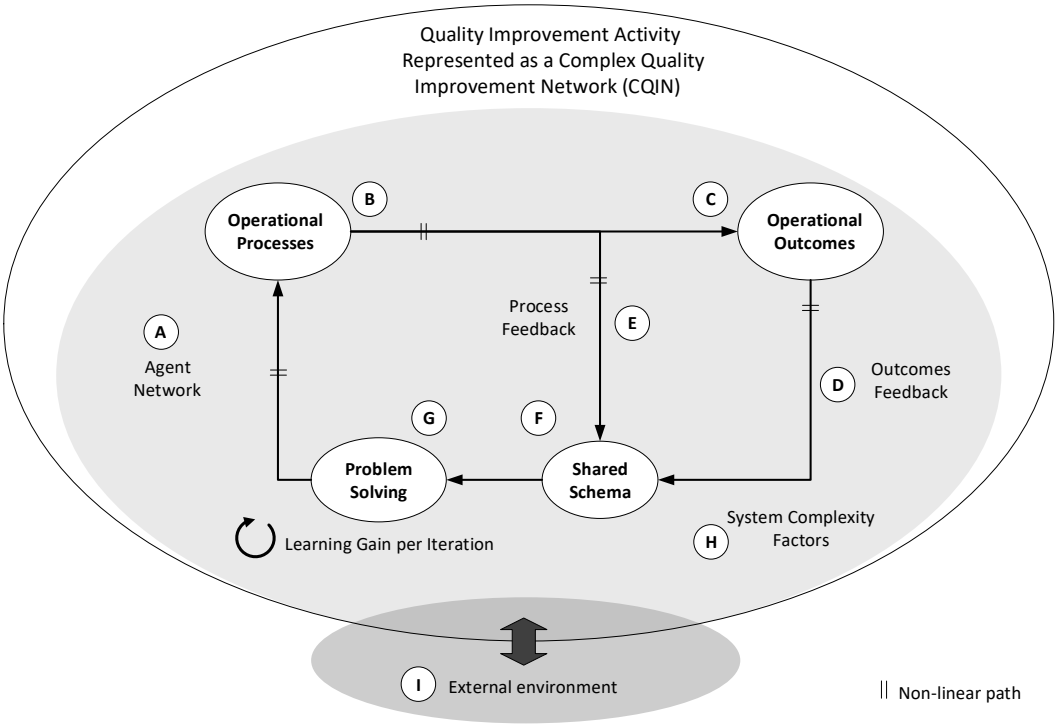


Figure 4.1: Generic Complex Quality Improvement Network (CQIN)

The model elements shown in Figure 4.1 are conceptual constructs. They are abstract variables *and* relationships that we are not able to directly observe (Meredith, 1993). Definitions for the CQIN constructs are presented in Table 4.1:

Table 4.1: CQIN Conceptual Constructs

CQIN Construct	**Definition**
Agent network (A)	The agent network is made up of all the participants and affected parties of the QI activity. For healthcare QI activity, the agent network encompasses clinicians, patients, QI practitioners and supporting functions, within the interpreted environmental boundaries.
Operational processes (B)	Operational processes and activity as relevant to the improvement context, for example – processes and applicable clinical protocols in a given healthcare context.
Operational outcomes (C)	Defined, measurable process and system outcomes.

CQIN Construct	Definition
System signals (D), (E)	Feedback on the health outcomes (i.e. changes in specified outcome measures) as well as wider system feedback (such as changes in the environment, resources, process timeliness or costs).
Shared schema (F)	The level of *coherently aligned* goals, assumptions, priorities, and improvement options across the agent network.
Problem solving (G)	The domain knowledge, problem-solving strategies, and resources (time, funding, cognitive bandwidth) available to the network.
System complexity factors (H)	The problem space is measured by the system complexity factors present. Using Wiesner and Ladyman's (2019) conditions for complexity, these are: (i) numerosity of elements; (ii) numerosity of interactions; (iii) diversity of elements; (iv) non-equilibrium; and (v) feedback. Products (outcomes) of complexity are: (i) non-linearity; (ii) self-organisation; (iii) adaptive behaviour, and (iv) memory (Wiesner & Ladyman, 2019).
External environment (I)	The external environment is that part of the contextual setting that falls outside the *interpreted and enacted boundary* (Nair & Reed-Tsochas, 2019). As this boundary is dynamic and may vary between agents, it acts as a moderating factor, in the form of *coevolution constraints* (impacts from shared and adjacent networks).
Learning gain	The iterative nature of CQIN is difficult to convey in a static picture. CQIN functions as a series of learning cycles. A learning gain from each iteration is posited as critical to acquiring new knowledge within the agent network (Deming, 1986; Doolittle, 2014; Reed et al., 2016).

4.7. The CQIN Conceptual Model as a Theory of Change

A fundamental tenet of CAS is the emergence of macro-level outcomes from a very limited set of agent rules or behaviours. To establish a genuinely CAS theory of change in socio-technical systems, demonstrating the emergence of complex outcomes from simple rules is necessary. The theory of change represented by the CQIN model is expressed via H4 and H5:

> H4: *Within socio-technical CAS, system-level change in the form of goal-directed improvement is achieved through coordinated updates and alignment of agent schema*.
>
> H5: *Updating and aligning agent schema is achieved within the constraints of a limited set of agent behavioural rules.*

Evidence to support H4 and H5 would show that simple agent behavioural rules for cooperation, learning, and updating internal schema, rather than being interpreted as simplifying or diminishing human agent activity, can instead be treated as CAS leverage points (Holland, 2012; Khan et al., 2018; Wilson & Kirman, 2016).

The conceptual model operationalises an underlying theory. Satisfactory theory requires: (i) definitions of the theory elements and constructs; (ii) identified domain boundaries for applicability; (iii) propositions for the relationships/associations between constructs, and (iv) the ability to make predictions or provide explanations (Wacker, 2008; Weber, 2012). Definitions for the CQIN model constructs and their interactions have been introduced. The boundaries of the model and the applicable domain for the theory (QI within socio-technical CAS) have also been specified. A high-level structure for the relationships has been proposed. Examining the relationships and interaction of the CQIN variables in detail, and providing explanations for QI within socio-technical CAS, is the objective of the subsequent chapters.

Part 2: Methods and Results

Chapter 5 Research Design

5.1. Introduction

This chapter begins with a discussion of the ontological, epistemological, and methodological considerations relevant to the research topic in section 4.2. Section 4.3 considers research methods that could potentially be used to investigate the research question. The challenges of empirical measurement of complex systems are discussed in section 4.4. The chosen research methods are then introduced and described in depth in section 4.5. This sub-section covers design, data collection and analysis. Within these categories, the use of simulation modelling, an applied case study, secondary data, thematic analysis, qualitative comparative analysis, and Bayesian network construction are all described. Section 4.6 sets out the steps taken to ensure ethical research conduct and the Section 4.7 concludes the chapter with a critical review examining the strengths and weaknesses of the chosen approach.

5.2. Ontological and Epistemological Perspective

Researchers have a responsibility to understand the different beliefs and values that influence how fundamental questions about knowledge are answered. Ontology is the philosophical domain behind the question 'what is real?' Epistemology explores 'how can we know this?' (Bryman, 2012; Guba, 1990). Answers to these questions determine how researchers attempt to understand a topic. The integration of ontological and epistemological views generates methodologies consistent with the chosen beliefs, covering the general approach and the specific data collection and analytical methods. These linked belief structures are referred to as paradigms (Guba, 1990). Relevant to this study is adequately positioning the use of quantitative and qualitative methods, because both methodologies are required.

Scholars of research methods acknowledge the philosophical and historical connections between research methods and paradigms but take care not to overstate them (Bryman, 2012; Tashakkori & Teddlie, 1998). Yin (2018) goes as far as to purposefully avoids the quantitative/qualitative descriptor when defining case research as a method. These views are relevant here given the very strong association with case studies as a method of qualitative research. Committing rigidly to an "embedded methods" argument, namely, the belief that methods inherit essential features of an underlying paradigm (Bryman, 2012), limits the methodological options for this research, which requires a

non-reductive analytical approach such as case study research. The following consideration of ontological and epistemological concerns is based on i) what the literature revels as to how other researchers have approached studies in this field; ii) the specific requirements of the topic, and iii) the research question and objectives.

Within the healthcare QI literature, there are consistent calls for theory-driven and generalisable approaches to QI research (Davidoff et al., 2015; Dixon-Woods & Martin, 2016; Khan et al., 2018; Kaplan et al., 2011; Walshe, 2007). A meaningful contribution to theory is an important consideration for this study. Researchers acknowledge that both quantitative and qualitative research can support theory development (Bryman, 2012), and this is also borne out in the literature. As two representative examples, Kreindler (2017) on the application of Lean within QI, and Manukyan et al. (2013) on the structure of agent networks, each highlight how successful research on this specific topic can be conducted within either a qualitative or a quantitative design.

CAS research straddles the natural science and social science domains (Holland, 2006, Kauffman, 2000; Miller at al., 2007). Consequently, research based on the natural science paradigm of a positivist approach is common, but a constructivist-interpretivist paradigm is not ruled out. Agent relationships, communication, and interpretation of the environment are essential to CAS (Holland, 2006, 2012; Miller et al., 2007). Within socio-technical CAS, human agency is a critical and fundamental component of the topic. Researchers exploring agent cooperation and shared understanding within a socio-technical CAS could legitimately argue a constructivist paradigm is essential to truly understand this interpreted context from the perspective of the agent. This point is fully acknowledged, but it does not rule out a mixed-method approach, or an objectivist ontological stance, namely, the belief that objects can exist independently of our awareness of them (Bryman, 2012).

Turning now to the research question, four hypotheses have been put forward to be measured by empirical data and deemed plausible or rejected given the evidence. Data collected will be a mix of quantitative (simulation experiments) and qualitative (case study) data, so the research is a form of mixed method research (Tashakkori & Teddlie, 1998). The approach uses a predominantly deductive mode of inquiry, with a starting theoretical model and ongoing hypothesis refinement based on data collected. Multiple modes of data collection are necessary for data and theory triangulation, verification of

simulation experiment results, and to gain more understanding of facets that are difficult to observe or measure. Despite the constraints, attributes of the system are being defined and observed, and these attributes can legitimately be said to exist independently of the system agents' or observers' understanding of them. A positivist epistemological standpoint, together with an objectivist ontological position, may well be appropriate for this research. However, the question of subjective, socially constructed reality should not be ignored. An example of this subjectivity is any perceived hierarchal relationships between agents, and the effect these have on cooperation. A further problem is the holistic nature of CAS, with the need to study, as much as possible, the full context that agents are operating within. Given the social science domain of this study, and the irreducible nature of CAS, a purely positivist paradigm is expected to be insufficient.

A better fit with all the requirements for this research is provided by critical realism. As a philosophical position that combines a realist ontology with a subjectivist epistemology, critical realism accepts that 'things' exist independently of perception or interpretation, but that the knowledge we can access about these things is theory-dependent and therefore subjective (Fryer, 2020). Critical realism focuses on causality but not the universal laws of a purely positivist paradigm, nor the subjective narrative of constructivism. Critical realism considers causal mechanisms as causal 'tendencies' and uses retroductive reasoning to explore cause and effect. Retroductive reasoning asks what is the best explanation for the observed data, given what we know? It seeks defendable explanation rather than prediction (Fryer, 2020; Gorski, 2013; Wynn & Williams, 2012). Within CAS research it is highly likely that the research will be observing the effects of underlying mechanisms not directly observable, but nevertheless explained by theory. The emergence of higher order effects is fundamental to CAS and also an intrinsic feature of critical realist thinking (Elder-Vass, 2005). Positioning a clinical microsystem undertaking QI as a social CAS supports a critical realist stance, by acknowledging the role of the agent network structure and interactions in any cause-and-effect analysis (Wynn & Williams, 2012). Critical realism has become more common in technological and CAS research (Preiser, Biggs, De Vos, & Folke, 2018; Tsang, 2014; Wynn & Williams, 2012). Given this very strong alignment with the nature of socio-technical CAS, a critical realist research paradigm has been adopted as the best fit for this study. The detailed design, data collection and analysis are coherently constructed to identify causal explanations consistent with a critical realist world view.

5.3. Research Methodology Options

The research question for this study is *What system factors influence the success or failure of healthcare QI activity within complex adaptive systems?* A sensible strategy for exploring such a broad research question is to follow a what/why/how questioning sequence (Blaikie, 2000). What is going on? Why is this happening? How are these outcomes achieved? For this research, the 'what' is the identification from the literature of inconsistent QI outcomes. The 'why' (also postulated within the literature) is the context of CAS. This study now asks a legitimate 'how' question in the form of *how can successful QI be achieved within CAS?* The conceptual model introduced in Chapter 4 provides the framework for answering this question. The CQIN model represents a theory of change within socio-technical CAS and the hypotheses implied by the model are testable. For example, the hypothesis that a high level of *shared schema* amongst networked agents positively influences QI success (Cronin & Weingart, 2007). The conceptual model also allows for the examination of counterfactuals, for example, the plausible argument that skill and fidelity to improvement methods may also be strongly influential.

5.4. Measuring Complex Systems

A significant barrier to building generalisable theories of complex social systems is that these systems are difficult to empirically measure in a way that allows for comparison across cases. The observable system behaviour evolves from countless interactions that agents have with each other and the environment, and it may be impossible to reliably observe them all. Compounding the basic issue of access, the environment (system) boundaries are open and highly dynamic. Even with willing participants, the measurement burden required to identify, for example, just the inter-agent communications, is formidable. The same constraints apply to validating the individual agent interpretations of their system or self-descriptions of their actions.

Unique starting conditions are also critical to CAS, and this makes exact cross-case comparison very difficult. The measurement challenge may explain in part the state of the applied CAS literature. High-level CAS principles are often referenced from the source theories and then used to describe new contexts. Academic researchers and practitioners appear to accept the validity of this transference, with few exceptions (as discussed in Chapter 3). Moving beyond the use of metaphors and descriptions to

proposing causal explanations supported by data from multiple cases will greatly assist improvement practitioners. This research manages the difficult challenge of measuring improvement within CAS in two ways:

1. Using generalisable simulation settings to examine the relationships and behaviour of critical CAS variables.
2. Providing a parsimonious conceptual framework to assess the widely diverse and/or incomplete language and reporting of published QI within CAS case studies.

The attributes of interest are abstract and often not recorded during conventional improvement activity. Understanding individual agent perceptions and motivations directly from the agents involved requires a very high burden of agent questioning, self-reporting, or observation, as well as the tracking of many simultaneous interactions. Case study research is an appropriate strategy for collecting rich qualitative data, such as an agent's perception and insider knowledge as to what they believe is happening. Case studies also limit reductionism, establishing wider context alongside theoretical constructs of interest. The entire case can become the unit of analysis, rather than isolated variables (Yin, 2018). Restricting the study to a very small number of cases would mitigate, but not eliminate, the measurement burden. A single or limited sample approach also raises generalisability concerns. Realistically, the selection of directly observed empirical cases is limited to those scenarios where the researcher can establish the relationships and cooperation necessary for an intrusive level of data recording. In other words, the case(s) would be a convenience sample, and some form of comparison or replication is required to make defendable claims of generalisability. The unique contexts of each CAS make cross-case analysis difficult, but not impossible, provided any generalisable elements are identified and available across multiple cases. Ideally therefore, multiple cases would be preferable, if a solution to the data collection problem can be found.

With reference to the methodological conventions of the subject areas, case research is a valid choice. Healthcare improvement research draws heavily on a shared history of quality theory and practice from non-healthcare contexts (Institute for Healthcare Improvement, 2023). As identified in section 3.2, this shared history with quality management also extends to complexity theory (Chandler et al., 2015; Khan et al., 2018).

Within the specific domain of operations and management research, Voss, Tsikritsis and Frohlich (2002) advocated case research as an appropriate methodology particularly for theory building, testing and refinement (Voss et al., 2002, p. 198). A contribution to successful theory building or refinement has been noted as a critical success factor for QI research. Eisenhardt and Graebner (2007) promote theory building from cases, pointing out that case study research (designed well) is a very effective bridge between "rich qualitative evidence" and deductive research (Eisenhardt & Graebner, 2007, p.25).

Several features of the research topic (QI within CAS) further support the use of case study research. For the most part, the QI activity being examined is contemporary activity, thus requiring field research (observation, interviews, and analysis). There are many variables that cannot be controlled, and the phenomena of interest and the context are not easily separated. This 'indivisibility' is perhaps the strongest argument for a research design that begins with the premise that any variable or contextual factor within the research scope is potentially of interest (Khan et al., 2018).

A final argument for selecting case research for this study is the objective of uncovering process as well as structural relationships (Mills et al., 2019). Quantitative measurement on its own is unlikely to reveal all the processes taking place between CAS agents, and qualitative data such as interviews and observation will be essential to understanding CAS activity. Case research can successfully incorporate this mixed method data collection and analysis (Bryman, 2012; Yin, 2018).

A survey research design is possible but severely limited by the reliance on the *a priori* specification of the survey questions and limiting the survey to a realistic size. Critically, it is already established in the literature that complexity concepts are not applied consistently nor well understood (Thompson et al., 2016). Variation in interpretation will limit the usefulness of the information gathered. Also problematic is the range of data required to adequately assess CAS behaviour. The research question is exploratory, in that it is not seeking to confirm or disconfirm one specific causal relationship. An ability to collect a wide range of data from observations and guided conversations is therefore necessary to complement the range of potential explanations inherent within the conceptual model (Yin, 2018).

Studying the history of a given healthcare QI CAS would also offer a useful perspective. However, a history research design relies on adequate and consistent

documentation, which is unlikely to exist or cover all the features of interest. This point has been made in the healthcare improvement reporting literature (Damschroder et al., 2009). A history research design would also require time for extensive interviews relating to past practices, introducing reliability issues, and the memory of participants. The lack of standardisation and reporting consistency also means that examining past activity will be very difficult. Given the high number of variables within CAS, it is unlikely that legitimate relationships between variables could be established from historical data alone.

The research topic is not suited to conventional experimental design. A clinical microsystem conducting deliberate improvement efforts contains a very large number of variables, which cannot be easily quantified or controlled within one experiment. The challenge is multiplied if attempting to control variables and reproduce results across multiple experiment runs. A quasi-experimental design may allow for more uncontrolled variables and the lack of random assignment, but the difficulty of managing experimental protocols remains (Bryman, 2012). Experiments require the identification of intervention and control groups. It is extremely unlikely that real world QI projects would allow this level of manipulation, deliberately failing by withholding a hypothesised intervention.

The important research design requirement can be summarised as follows. This research requires a design that can resolve the conflict between manageable data collection (using small or intermediate sized samples) and the ability to explore the theoretical constructs in depth (requiring more examples or cases and the opportunity to collect a lot of data).

5.5. Research Method

5.5.1. Overview of Research Design

Figure 5.1 is an overview of the research design, showing the steps in the research design and the rationale for each step. The expected contribution resulting from the research is also included:

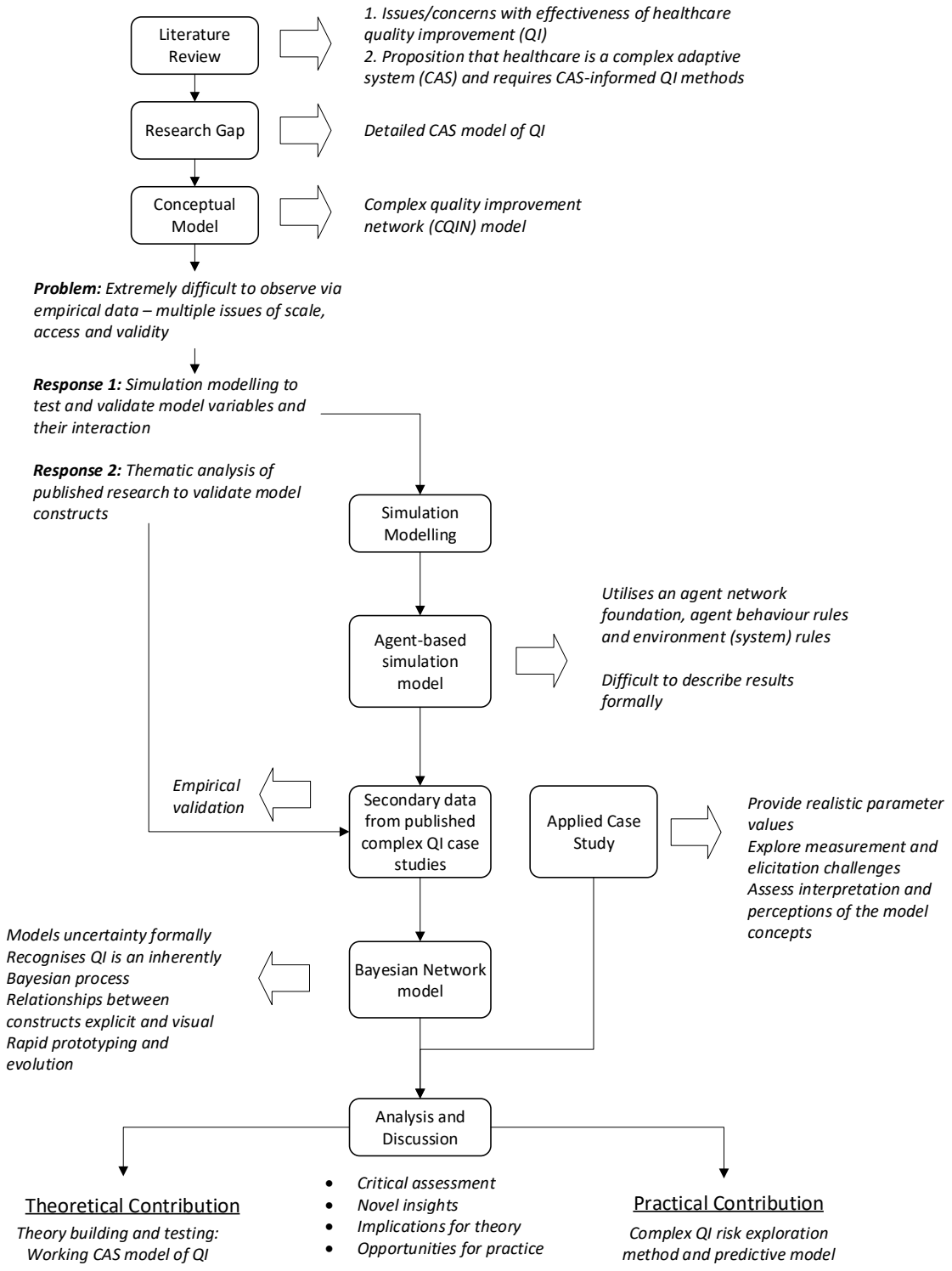


Figure 5.1: Outline of Research Method and Expected Contribution

The research follows a mixed methods design, using simulation experiments to establish the initial plausibility of the theory, and thematic analysis and QCA to then examine the evidence for the model. An applied case study provides initial face validity for the simulation model and qualitative insights in practitioner understanding and application of

CQIN. These observations about practitioner interpretation are not available via the secondary data analysis. Although perhaps not strictly necessary within a simple 'model then validate' research design, the representation as a Bayesian network allows formal description and easy visualisation of the underlying model. Additionally, if the underlying theoretical model appears valid, it is a logical and straightforward next step to build a simple predictive model. The Bayesian network achieves this by combining the refined theory with the empirical validation data. The individual research components form a coherent whole relevant to the research topic. Collectively, the methods within this research design navigate the measurement challenge of complex systems, in terms of the infeasibility of direct observation and the gaps in recorded information.

In summary, the research design seeks to answer the research question and objectives via five research components as shown in Figure 5.2. The conceptual model is operationalised as an agent-based simulation model, and the secondary data is used to validate the model. Are the model constructs and simulated behaviour supported by empirical evidence? The applied case study provides realistic starting parameters for simulation, and access to practitioner interpretations of CQIN. Finally, a predictive model is developed by bringing all the research streams together. The research mode is primarily deductive, testing and refining the implied theory of the conceptual model via multiple modes of enquiry, and examining the empirical evidence to validate the model. To refine the theory, relevant CAS principles are identified, applied and then evaluated within the defined socio-technical domain. The deductive approach contrasts with the inductive theory-building techniques commonly associated with case study analysis (Yin, 2018). However, the notion of theory *refinement* is key, with each research stream providing insight and refinements to iteratively update the original conceptual model. The refined CQIN model and theory is discussed in section 11.6.

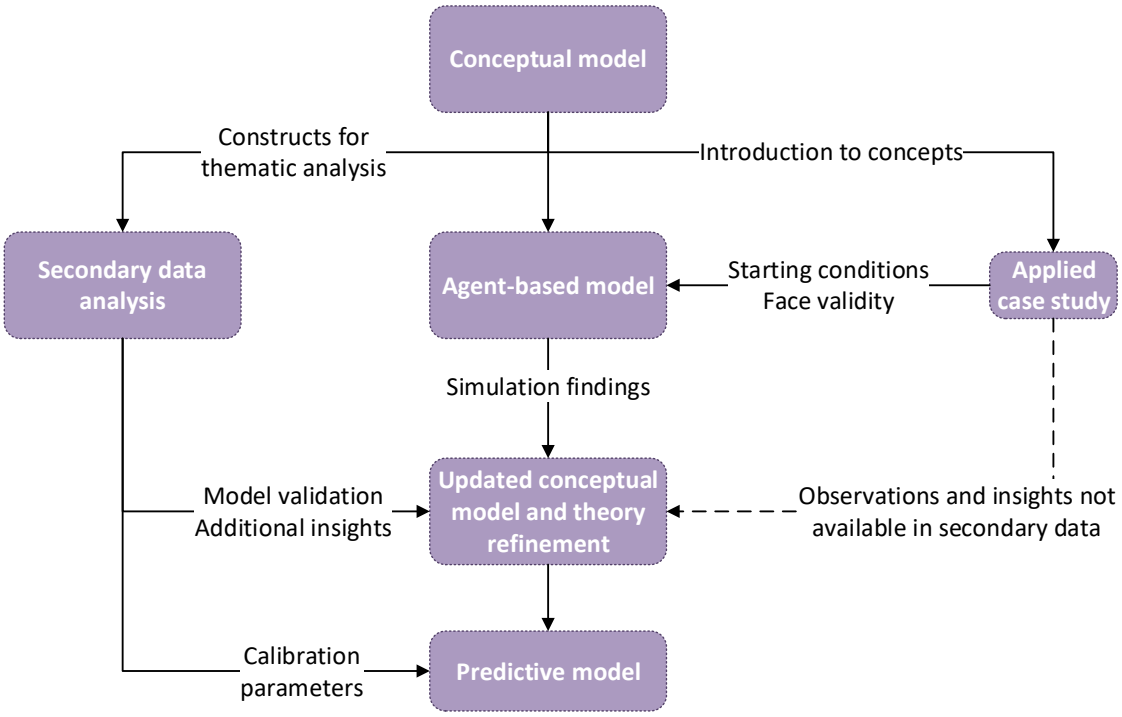


Figure 5.2: Integration of Research Components

5.5.2. Simulation Modelling

Simulation methods allow in-depth experimentation without the practical and ethical difficulties involved with manipulating real-world QI activity. Simulation is also useful for developing theory. The road map for developing theory via simulation methods proposed by Davis, Eisenhardt and Bingham (2007) guided the simulation approach for this study. The key steps of the roadmap are reproduced in Table 5.1, mapped to the application of each step within the study.

Table 5.1: Roadmap for Developing Theory Using Simulation Methods

Step	**How this step was implemented in the research**
Begin with a research question	*What system factors influence the success or failure of healthcare QI activity within complex adaptive systems?*
Identify simple theory	CQIN conceptual model
Choose a simulation approach	Agent-based modelling identified as the best alignment with CAS

Step	How this step was implemented in the research
Create computational representation	CQIN agent-based model constructed in NetLogo modelling software
Verify computational representation	Established face validity using a real-world scenario. Test the behaviour of the model under full range of conditions. Replication
Experiment to build novel theory	Three experimental designs: 1. QI project simulation 2. Influence of Schema share effectiveness variable 3. Sensitivity testing of all CQIN variables.
Validate with empirical data	Use of thematic analysis and qualitative comparative analysis to see if the posited constructs and behaviour are reflected in relevant empirical activity.

(Adapted from Davis, Eisenhardt & Bingham, 2007)

The starting point in the roadmap is identifying an appropriate research question that can be addressed by what is described as 'simple' theory. Simple theory is not intended to be universal theory, nor is it fully developed. Simple theory should have sufficient basic constructs in place to enable exploration, but still require more detailed understanding and refinement that simulation can provide (Davis, Eisenhardt & Bingham, 2007). The research question and CQIN conceptual model, representing the simple theory, have been introduced.

Next, a simulation approach should be determined, and the computational model created. The computational model must be verified to ensure it has been constructed correctly and functions in line with the intended specification. Following basic functional tests, face validity of the CQIN ABM simulation was conducted using an empirical case study.

5.5.3. Face Validity Case Study

One in-depth empirical case study was identified to provide plausible operating parameters and face validity for the ABM simulation. Face validity is an initial, non-scientifically rigorous assessment that the model is suitable for its intended purpose (Fitzner, 2007). This study also provided an opportunity to test clinician and QI practitioner interpretation of the abstract CQIN constructs. The insights from practical engagement with CQIN are presented in Chapter 10.

The case study organisation is a large private provider of aged residential care (ARC) services in New Zealand and Australia. This organisation maintains an interdisciplinary medications advisory committee (MAC) that oversees performance and improvement of medication-related activities. The committee covers clinical and operational functions and includes general practice representatives, a pharmacist, geriatrician, specialist researcher, nurses, and improvement support staff. A major on-going QI programme run by the MAC is minimising antipsychotic prescribing for patients with dementia.

Seeking to minimise potentially clinically unnecessary prescribing involves a large heterogenous network of agents across multiple organisations and professional disciplines. Alternative models of care are required to reduce usage of antipsychotic medications. The number of agents involved, the number of parallel interactions, the requirement to align goals, assumptions, and understanding of progress all combine to present an example of complex problem solving.

Returning to the Davis et al. (2007) roadmap, following internal and face validity assessment, the simulation model was then ready for experimental design, with controlled interventions to test hypotheses and explore system behaviour under defined conditions. Chapter 6 contains the detailed description of the agent-based model. The last step in the roadmap is to validate the simulation findings with empirical data. Are the simulation findings realistic? What empirical data support the findings? To assess empirical validation, secondary case data from published cases in the healthcare QI literature was used.

5.5.4. Validation Using Secondary Case Data

Ideally, a comparative analysis of QI activity focusing on multiple positive and negative examples would be used to provide an optimal path for theory building and validation. In practice, limitations are imposed by using secondary data, even when data collection is

confined to published case studies in peer-reviewed academic journal articles. All information is self-reported, and subjective. Despite the availability of standardised QI reporting guidelines such as SQUIRE, many QI articles still have limited reporting of important QI information and concepts, and considerations of complexity concepts are highly variable. Most seriously, a very strong publication bias is present, where QI articles are overwhelmingly biased to positive reporting. 97% of the secondary data cases reported positive (successful) outcomes. These are important limitations, but once fully aware of them and their implications, appropriate analysis was still possible.

5.5.5. Measurement Techniques and Associated Analysis

The scoping review presented in Chapter 3 details the full data collection process for the secondary case study data. Thematic Analysis of this data was carried out using the defined CQIN constructs as the thematic framework. Thematic analysis is a qualitative analysis method that identifies patterns within narrative data. The patterns may be manifest (directly observable), or latent, where the matching criteria is only indirectly referenced or implied. Agreed criteria are required to determine what data meet the definition of a theme (Joffe, 2011). The defined CQIN constructs functioned as the thematic criteria for the analysis. Therefore, this study used a deductive approach to thematic analysis, examining the evidence for the proposed theory. Table 5.2 contains two initial examples to illustrate how the CQIN constructs were identified as themes within the text.

Table 5.2: Thematic Analysis Example

CQIN Construct (Theme)	**Text [Case ID]**	**Manifest/Latent**
Autonomy	*Clinicians know what is wrong with the system and are well placed to test solutions if allowed to do so [42]*	Manifest
Interpersonal Factors	*However, in the single-minded pursuit of improvement science, we suggest that something gets lost - the appreciation of humanity, the art of conversation, the ability of leaders and staff to listen to each other, to explore their experience, to voice their concerns and hope [60]*	Latent

Joffe (2011) highlights the danger of proceeding only in a deductive manner, looking for a pre-determined theme, and ignoring other potentially relevant information within the data. Additional themes, not included in the conceptual model but relevant to the context of QI and CAS, were also identified and described.

Crisp Set Qualitative Comparative Analysis (csQCA, hereafter QCA) was then conducted on the secondary data themes. QCA is a technique that allows variable-level comparison across cases whilst still preserving the whole case data. QCA thus has elements of qualitative (case-oriented) and quantitative (variable-oriented) methods. QCA explores causality via 'multiple conjunctural causation' (Berg-Schlosser, Meur, Rihoux, & Ragin, 2009; Ragin, 1999). This is achieved by establishing sets of variables associated with a designated outcome variable, within each case. Multiple sets or 'recipes' for successful QI are compiled. Coverage (within case) and consistency (across cases) was assessed. Boolean minimisation is applied to identify minimum solutions across multiple sets. Conventionally, *sufficient* conditions (present in positive outcomes) and *necessary* conditions (must *always* be present in positive outcomes) are identified (Rihoux & De Meur, 2009). A critical distinction noted for this research was that there were insufficient negative cases to determine necessary conditions — those conditions

that must *always* be present in *all* cases with a successful outcome. The identified sets of improvement conditions therefore represent *sufficient* conditions only (i.e. those sets of conditions which have been reported to result in successful QI).

The final analytical strategy applied was to interpret the CQIN conceptual model as a Bayesian Network (BN). An acknowledged limitation of ABM is that the models can be difficult to share and interpret formally (Wilensky & Rand, 2015). Analysing the CQIN model as a BN enables further analytical possibilities and is also appropriate for the general subject matter themes of problem solving and complexity. BNs are probabilistic models comprised of a directed acyclic graph showing parent/child relationships between entities and probability tables associated with each node. BNs are a practical tool for prediction and causal inference, especially suited to complex system analysis because they can cope with incomplete observations as well as sparse and noisy data. The relationships between system components are expressed formally and transparently using joint and conditional probability. BNs can learn via imported data or expert reasoning and can be iteratively updated as more information becomes available. In short, they are an approach well suited to uncertainty and iterative learning (Constantinou & Fenton, 2018; Fenton & Neil, 2019). In their scoping review of BNs developed for clinical decision support, Kyrimi, McLachlan, Dube, Neves, Fahmi & Fenton (2021) raise cautions to the enthusiastic exploration of BNs. They find important gaps in the reporting of the clinical purpose; the network development process; articulation of the network structure and parameters; expert knowledge elicitation; learning algorithms; model validation, and subsequent use of the model in practice. These cautions formed a good practice checklist for the development of the CQIN BN model.

Bayesian inference, updating a hypothesis with new evidence, is the foundation for QI. A plan-do-study-act cycle starts with one level of information that is updated when new evidence is obtained. When there is high level of complexity present, incomplete knowledge is virtually guaranteed. A BN allows the incomplete knowledge within a CAS to be estimated and quantified. Creating the CQIN BN model added 5 analytical contributions: i) A second instance of the CQIN constructs, demonstrating that the concepts can be reconstructed in different modelling methods; ii) a more accessible and visual representation of the variable relationships; iii) formal description of the variable relationships; iv) exploration of empirical validation using the QCA analysis to update the node probability tables, and v) a risk assessment tool for improvement practitioners.

5.6. Ethical Considerations

The Massey University Code of Responsible Research Conduct (2015), Code of Ethical Conduct for Research, Teaching and Evaluations Involving Human Participants (2017), and a summary of the Te Ara Tika Guidelines for Māori research ethics were reviewed by the author (Massey University, 2015, 2017). Ethical considerations for this project were discussed with research supervisors and the management contact for the empirical case organisation.

The fundamental ethical criteria of voluntary participation, informed consent, maintenance of privacy and confidentiality, respect for participants and the sharing of results (Burns, 2000), were directly relevant to the applied case study component of the research and addressed via the communication plan and management of the project. The participation of busy clinicians demands due respect for their time and the author was responsible for ensuring data collection activity was well planned and made the most of everyone's time. An information sheet, consent form and practical guide to the abstract language of the conceptual model were developed to facilitate data collection. Interaction with participants was first organised with the clinical or management lead for the case study organisation. To encourage open and frank responses from the applied case study participants, confidentiality was maintained, and symmetry of outcomes was emphasised (i.e. respondents were supported to contribute their opinion anonymously, and also encouraged to say if something was not working as planned).

No patient identification or clinical information was accessed or analysed in this research. The research thus falls into the general category of "audit and related activity" as defined by the (Ministry of Health) Health and Disability Ethics Committees. More specifically, the research fits the definition of programme evaluation/evaluation studies – examining structure, process and outcome of selected activities. The lack of identifiable patient data or clinical information, within the context of an evaluation study, means that this research is considered out of scope for the Health and Disability Ethics Committees review (Health and Disability Ethics Committees, 2018). These same factors positioned the research as low risk within the Massey University Code of Ethical Conduct for Research, Teaching and Evaluations Involving Human Participants (Massey University, 2017). Massey University Human Ethics Low risk assessment 4000023583 was issued for the study in October 2020.

To prevent any conflict of interest, no improvement initiatives which involved the author were included. Finally, data collected during the research has been maintained in a secure database and is available to reviewers and other researchers as appropriate. The data repository has been backed up to prevent single point of failure.

5.7. Critical review of the research method

It is necessary to briefly outline why three established and potentially relevant analytical lenses were *not* adopted. First, *Actor-Network Theory*, with its focus on networks, relationships between system actors, and the dynamic creation and evolution of these relationships, is one seemingly obvious candidate. The ontological basis of Actor-Network theory appears to be contested, but the reason for not exploring it further was one of method. This approach requires deep immersion and access to the relevant interactions within a given network (Creswell, Worth & Sheikh, 2010). The close and detailed observation of system 'actors' required by this framework was simply not possible. The second framework, *Organisational Cybernetics*, including *Viable Systems Modelling*, has a good overlap with the subject matter and contemporary application to the topic (Akmal et al., 2021). For this study, the author believed that a more general starting point was required to establish the CAS principles — a generalised system, rather than the special case of an organisation perceived as a system. A third analytical approach considered was *Game Theory*, specifically evolutionary game theory. With the ability to represent and explore cooperation amongst groups of agents, this is a logical framework to consider. The difficulty with applying game theory to the current study is adequately accounting for the intentionally shared understanding and assumptions behind any cooperation. Interactive (social) decisions require more than utility maximisation, for example, interpersonal factors. Coleman, (2003) argues that instrumental rationality and the common knowledge assumptions of game theory do not reflect the empirical results of cooperative game studies. Nevertheless, as game theory continues to evolve, game theoretic explorations are a promising avenue of research for cooperation amongst a network of adaptive agents (Coleman, 2003; Scatà et al., 2016).

The twin disruptions of COVID-19 and New Zealand healthcare system reforms had an impact on the early stages of this research. The relevant impact was not personally on the author, but on the availability and capacity of clinicians and healthcare improvement practitioners. Under the intense pressure of these disruptions, opportunities for engagement with a non-clinical research project were extremely limited. This situation

required a design that minimised disruption to working clinicians. The design and methods for this study nevertheless follow an organising logic relevant to the research question and the nature of CAS. The research objectives were achieved via the articulation, development, and refinement of the underlying theory. Simulation experiments established that the minimum CQIN constructs, and the hypothesised interactions were plausible. To validate the model and simulated findings, secondary data from published QI cases referencing complexity were analysed. Thematic Analysis and QCA were fully appropriate for this qualitative analysis task. QCA is increasingly being used to investigate complex system causality. A review on the use of QCA in public health interventions found that QCA techniques allow investigation of complex causality within cases studies in a systematic manner. Notably, this review also highlighted the lack of data related to *ineffective* interventions as a weakness (Hanckel, Pettigrew, Thomas & Green, 2021).

Although the method for the current study was predominantly deductive in mode, inductive analysis also occurred in both the simulation findings and the thematic analysis, refining the theory in an iterative manner. An argument can be made that starting with an external theory (CAS) and applying it in a deductive approach adds useful pluralism to the research on complex healthcare QI, as inductive theory-building dominates the field. Reed et al. (2018) and Sharma et al. (2024), are recent exemplars of the dominant approach. Finally, the conceptual model and the QCA findings were used to construct a predictive model. The Bayesian network bringing the theory, experimental findings, and empirical support for CQIN together. The mature iteration of the CQIN BN model is ready for ongoing empirical evaluation.

A level of subjectivity within the thematic analysis is acknowledged because the interpretation and coding activity was limited to the author. This is a limitation of solo research not part of a wider research laboratory or programme. The use of secondary data also means that interpretations of QI, complexity, and reporting of factors influencing success were subjective to the researchers describing their results. Countering this limitation, using published articles from peer reviewed journals provides a base level of quality assurance. QI cases studies are published so that other researchers can learn from them and generate new ideas. Existing cases studies should be a legitimate source of information to support theory development. Clearly defined criteria, informed by theory, supported consistent pattern recognition across the cases.

Chapter 6 Development of the Agent-Based Model

6.1. Introduction

This chapter describes the development of the ABM in detail. General requirements for developing an agent-based model are first set out in section 6.2. The detailed construction of the CQIN simulation model then follows in section 6.3. Section 6.4 outlines the abstraction and simplification assumptions. The calculations behind the agent behaviour are presented in section 6.5. Section 6.6 then provides a detailed description of the model behaviour with representative examples. Verification and validation are discussed in section 6.7. This section first introduces the specific factors for testing ABM and then describes the tests associated with confirming the model construction. To conclude the chapter, section 6.8 summarises the essential features of the CQIN agent-based model.

6.2. Model Requirements

Computer modelling is perhaps the most common method of enquiry into CAS (Holland, 2006; Mitchell 2009). Computer simulations allow 'what if' scenario analysis and extend the evaluation of propositions into formal hypothesis testing via manipulations of model factors simply not possible using real-world observation. Agent-based models are recognised as the appropriate modelling approach for complex systems involving autonomous agents (Preiser, Biggs, De Vos, & Folke, 2018). Several mathematical foundations for agent-based modelling have been established, with the computation usually built from a *Finite Dynamical Systems* modelling archetype. The minimum structure and functions required for agent-based models built as a Finite Dynamical System are: an agent dependency graph; a rules-based update function, and an update mode (sequential or parallel updates). Deterministic and stochastic processes can be introduced into the variables and update procedures (Borrill & Tesfatsion, 2010; Laubenbacher, Jarrah, Mortveit, & Ravi, 2007). Several accessible agent-based modelling applications exist, including open-source systems such as NetLogo (Thiele, 2014; Wilensky, 1999). The CQIN simulation model was built using the NetLogo modelling software (Wilensky, 1999). NetLogo is regarded as a modelling platform suitable for researchers who lack the extensive coding skills required of other ABM software options and is widely used in the teaching of agent-based modelling fundamentals (Railsback, Lytinen, & Jackson, 2006). Despite (or perhaps because of) its relative ease of use,

NetLogo has been used for a substantial body of published ABM research in many knowledge domains (NetLogo, 2023a).

6.3. CQIN ABM Model Development

To be fully operationalised as an ABM, the CQIN conceptual model was extended with additional variables to enact the spreading process between agents, set value ranges and stopping criteria, and a system observation and reporting function (Laubenbacher et al., 2007). Table 6.1 presents the CQIN conceptual model now mapped to the specific, configurable parameters required to run the simulation model. The parentheses in the CQIN constructs column tie the operational model back to the generic model as previously introduced in Figure 4.1. The parameters involved in the threshold calculations are also indicated. The threshold calculations drive the model behaviour in accordance with the identified CAS principles (Section 6.5).

Table 6.1: CQIN Elements Mapped to Modelling Parameters.

CQIN Constructs	**Configurable Parameters in ABM**
Agent network (A)	● Number of networked agents (nodes)** ● Network type (preferential attachment or random) ● Minimum node degree ● Network centrality measures ● Proportion of network cooperating
Operational processes and Outcomes (B), (C)	● Outcome measure at start ● Targeted outcome measure value ● Elapsed time (time "units")
System signals (D), (E)	● The range of signal detection amongst agents ● Signal amplitude ● Signal validity (relevant to objectives)

CQIN Constructs	Configurable Parameters in ABM
Shared schema (F)	• Schema share effectiveness**
Problem solving (G)	• Problem-solving effectiveness**
System complexity factors (H)	• Problem space complexity** • Network rejection scope** • Network rejection probability** • Minimum cooperation threshold
External environment (I)	• External (coevolution) constraints** • QI phase**
Learning	• Learning gain per cycle** • Improvement increments • Number of planned iterations • Number of iterations completed (count)**

** Parameters involved in the threshold equations

One problem to be solved in the development of the model was how to incorporate the success or failure of any given agent interactions when these cannot be determined *a priori* nor observed. The CQIN model simulates the uncertainty associated with agent interactions by applying a stochastic procedure to each agent interaction. Setting a probability to each agent interaction overcomes the observation difficulty of not knowing which interactions will succeed and which will fail by randomizing this probability across all the agent network activity, with configurable parameters to adjust the probability of success.

The interaction of the identified parameters, via the threshold procedures, is essential to the system behaviour and the structural relationships and interaction of the variables is now examined more closely. Figure 6.1 presents the constructs from the CQIN conceptual

model arranged in a causal path diagram showing the hypothesised parent-child relationships and other shared dependencies between variables.

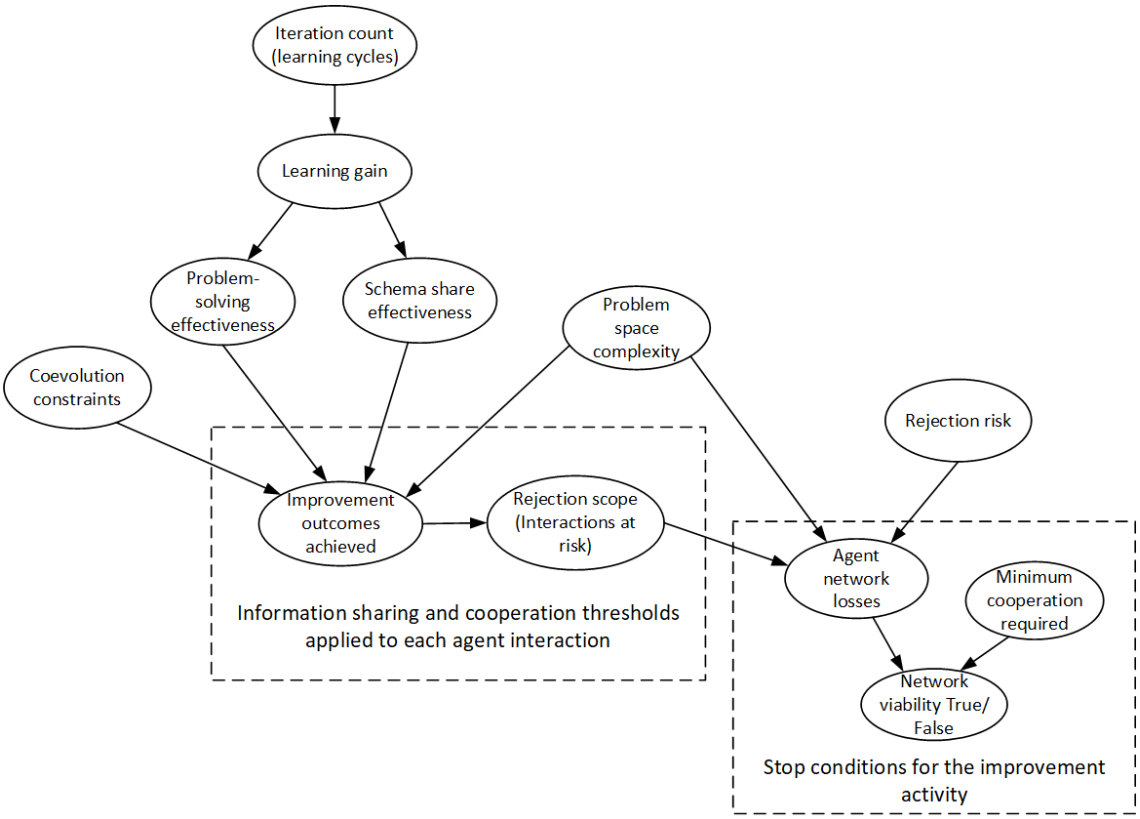


Figure 6.1: Hypothesised ABM Variable Relationships

Schema share effectiveness (the shared understanding between agents) was hypothesised as the dominant variable as it is a necessary pre-condition for problem solving. (Cronin and Weingart 2019; Dooley 1997; Holland 2006, 2012). Figure 6.1 shows how the *schema share effectiveness* has a direct causal influence on improvement outcomes alongside *problem-solving effectiveness. Problem space complexity* and *coevolution* (external) constraints have a moderating effect that can range from very little influence to fully inhibiting. The ability to learn from each improvement iteration positively influences S*chema share effectiveness* and *problem-solving effectiveness* to increase the level of improvement achieved. The notion of non-cooperation amongst agents was also introduced, using a *susceptible-infected-resistant* (SIR) virus-spreading model as the example. The relevant assumption is that unsuccessful problem-solving places ongoing agent cooperation at risk. In some scenarios, agents may eventually become 'resistant' to change. As the level of problem complexity and endogenous risk factors such as external shocks to the network increase, the potential for non-cooperation

also increases. At some threshold, it is possible for the agent network to fall below a minimum level of viability.

6.4. Model Assumptions:

The following assumptions explain the update process in relation to quality improvement activity and the intended level of abstraction for the agent interactions.

1. One iteration of the update sequence represents one full learning cycle – equivalent to a *plan-do-check-act* improvement cycle with an initial signal detection step added at the start.
2. The threshold calculations are intended to be sufficiently flexible to mimic the learning cycle phases, for example the *coevolution* constraint can be re-weighted to be active (or more strongly influential) in the "do" and "act" phases, which require more interaction at the network boundaries.
3. "Sharing information" and "cooperating" are abstract constructs and treated as equivalent in terms of representing agent interaction. However, they are functionally distinguished based on which learning phase applies at the time. The interaction varies depending on the improvement phase, for example: *notice this signal* (review phase), *is this signal important?* (check phase), *what should we do?* (plan phase), *can you assist us with this action?* (do phase). The coding of the threshold equations reflects the difference between sharing information and cooperating, with an assumption that the "do" phase is more heavily constrained by coevolution with related networks.
4. Shared Schema should be understood as shared understanding or aligned understanding; it does not mean identical views are held by each agent.
5. Similarly, "rejection" or "non-cooperation" are also abstract and may represent any of *unwilling or unable to receive, understand, agree or participate*. Any of these may apply in practice, but all have the same effect of inhibiting cooperation at the given point in time.

6. Agents represent any role within the clinical microsystems including patients and supporting teams. The wide variety of interactions between professional roles and communities is approximated by the preferential attachment network structure and the randomisation applied to any one interaction. Preferential attachment creates realistic social networks, for example collections of clinicians interacting

closely with each other and less closely with more distant communities in the network. Randomisation in the threshold equation mimics the unobservable and potentially inexplicable interactions where, for some reason, two closely connected individuals fail to cooperate.

7. Rejection /non-cooperation can be treated as permanent or able to be reset after each iteration. For this study the rejection level was reset after each full PDSA cycle, premised on the idea that the network could regroup after each completed cycle.

6.5. Threshold Equations

Within CAS, a small set of internalised agent rules are sufficient to produce complex agent activity (Holland, 2012; Kauffman, 2000; Mitchell, 2009). Agent behaviour in the CQIN model is governed by five rules only: (i) request to share information; (ii) request for cooperation; (iii) reject a request; (iv) accept a request, and (v) withdraw from the network. These five agent behaviour rules can be described using conditional probability statements.

Achieving a shared problem understanding and response across the agent network is conditional on: (i) successful transmission of information; (ii) acceptance of the new information; and (iii) ongoing cooperation of the receiving agents. The overall probability of achieving *intentional, system-level improvement outcomes* (*S*) across the agent network is shown in equation 1.

Equation 1:

$$P(S) = P(A|B,C) = \frac{P(B \cap C|A)P(A)}{P(B \cap C)}$$

Where:

A = the initial level of shared schema between agents

B = successful sharing of relevant information

C = accepting and acting on the relevant information

To establish successful information circulation across the network (*B*):

Equation 2:

$$P(B) = \left(\sum_{i=1}^{n} a_i \, P \left(x \; < DE^F \left(\frac{GE^F}{H} \right) \left(\frac{(100 - I)}{100} \right) \right) \right)$$

Where:

a = an agent attempting to share information

D = the schema share effectiveness;

E = the learning gain per improvement iteration;

F = the improvement iteration count;

G = the problem-solving effectiveness;

H = the problem space complexity; and

I = the coevolution constraints

$x = random\, x \; \in \mathbb{N}, x \; \in [1, 100]$

Equation 2 introduces the threshold calculations coded in the ABM. For each agent attempting to share with a connected neighbour, the interaction product of agent sharing/cooperation activity is compared to a random number between 1 and 100 (uniform distribution). The stronger the information-sharing and cooperation is, the greater the value of this product, and the increased probability of successful transmission amongst agents.

6.5.1. Rationale for the Interaction Product

The request to share/cooperate begins with an initial value of the *schema share effectiveness*, bounded between 0 and 1. This value is multiplied by a learning-gain, raised to the power of the number of improvement iterations. There is support for the idea that learning from practice increases following a power law distribution in the cognitive psychology literature (Anderson, 1993, 2005). For example, a 10% gain over 6 iterations would become 1.1^6 = 1.77. This combined value is then multiplied by the ratio of

problem-solving effectiveness over the *problem space complexity.* The 'effectiveness' of the problem solving is thus measured directly against the problem space being explored, including the aspects of complexity present. Problem solving also benefits from the learning gain. Finally, the interaction value is multiplied by the complement of the coevolution constraints (1- *coevolution-constraints).*

In summary, maximum problem solving within a given problem space, together with zero external constraints, will mean there is no degradation of the schema share value. Anything less than 1 for the problem solving and external constraint ratios will mean the schema share effectiveness diminishes proportionately.

The probability of accepting new information and continuing cooperation (*C*) is shown in equation 3:

Equation 3[4]:

$$P(C) = \left(\sum_{i=1}^{n} b_i \frac{P\big((x < J) \cap P(x < K)\big)P(x < K)}{P(x < J)} \right)$$

Where:

b = the agent receiving/acting on the new information

J = the range of network interactions at risk

K = the rejection probability

$x = random\ x \in \mathbb{N}, x \in [1, 100]$

In parallel with agents attempting to share/cooperate, their willingness to continue doing so is tested. Agent *b* might accept a request from agent *a* on the first attempt or the 2nd, 3rd, or *n*th attempt. At the same time, agent *b* may also reject all requests and become non-cooperative. A random number between 1 and 100 (uniform distribution) is compared to the joint probability of the number of agents at risk of rejection and the probability of rejection. A negative moderating influence from the problem space

[4] Equation 3 has been updated from the version included in Wilson, McLachlan, Dube, Potter, & Jayamaha (2023). A formatting error omitted the final term in the numerator (the prior). The text description is unchanged.

complexity (H) on the rejection risk can also be set. Higher risk levels increase the likelihood of rejection/non-cooperation.

6.6. Detailed Description of the NetLogo CQIN Model

The starting idea for the CQIN spreading process is a *susceptible-infected-resistant* (SIR) process used to model the spread of a computer virus throughout a computer network. The base 'virus on a network' model used is freely available from the NetLogo modelling commons (Stonedahl & Wilensky, 2008). Two essential ideas were retained from this original model. Firstly, the underlying stochastic spreading process achieved by comparing a value of interest with a random number. Second, the concept of an acceptance/rejection test running in parallel so that the agent state could be determined at each update step. The original code from Stonedahl & Wilensky (2008) for these two building blocks is shown in Figure 6.2

```
;; to spread-virus
;;  ask turtles with [infected?]
;;    [ ask link-neighbors with [not resistant?]
;;       [ if random-float 100 < virus-spread-chance
;;            [ become-infected ] ] ]
;; end

;; to do-virus-checks
;;  ask turtles with [infected? and virus-check-timer = 0]
;;  [
;;    if random 100 < recovery-chance
;;    [
;;      ifelse random 100 < gain-resistance-chance
;;        [ become-resistant ]
;;        [ become-susceptible ]
;;    ]
;;  ]
;; end
```

Figure 6.2: Spread and Schema-Check Code Examples (From Stonedahl & Wilensky, 2008)

6.6.1. Initialisation

For CQIN, rather than a computer virus, it is information and requests for cooperation that are shared amongst linked neighbours. The operation of the ABM is now described in detail with illustrations from the NetLogo editing interface. NetLogo uses a library of modular high-level code objects such as sliders, switches, input cells and reporting charts to build the editing environment.

The setup routine reads in the chosen starting parameters for the CQIN variables, as well as the improvement parameters of planned iterations and improvement gain per iteration (representing an average across the number of cycles). The agent network is

initially defined by the number of nodes and the minimum node degree for the network. An existing agent network can be imported using a *graphml* file format or generated automatically. The import functionality was developed and tested using a small fictional project, but not further applied for this study. As originally conceived, the specific structural features of an imported agent network can be examined within the model, but the challenge of accurately identifying the network nodes and edges must first be overcome. Network generation for a given number of agents can have random node connections or follow the preferential attachment convention consistently observed in social networks (Barabási, 2016). Links between agents can be set as directed (one way) or undirected (bi-directional). Once generated, the agent network is laid out to fill the visual frame. All agents are initialised in a 'susceptible' state, open to signals from the environment. The 'trigger' function updates a configurable number of agents with a new environmental signal. Figure 6.3 illustrates a network of 255 agents with 20 agents in an updated state (red), ready to attempt to share this new information with their connected neighbours.

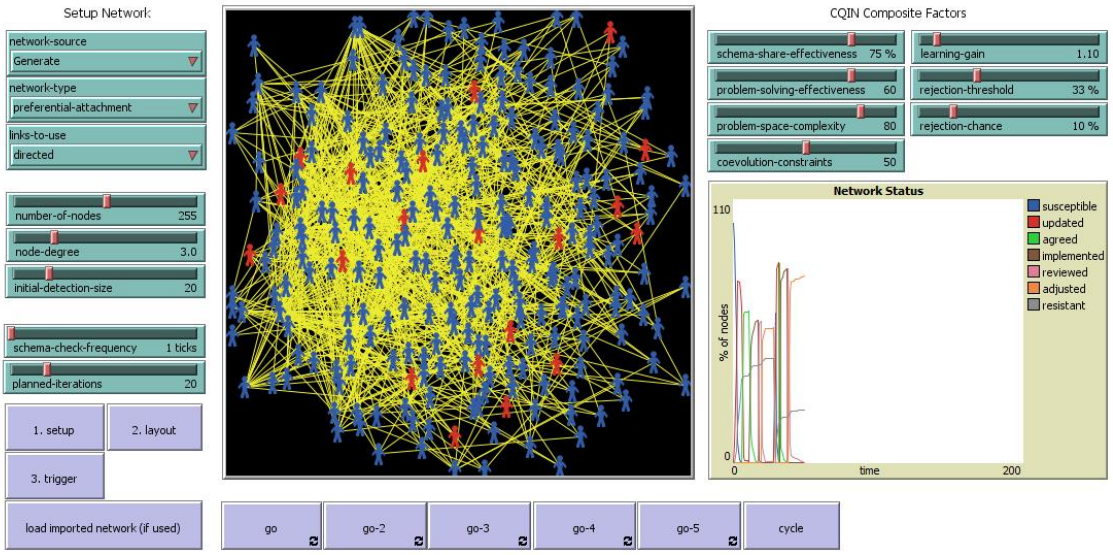


Figure 6.3: CQIN Editing Interface Example 1

6.6.2. Improvement Phases

Each agent with the updated environmental signal then seeks to share the update with their connected neighbours. Sharing follows a transmit and receive structure, where the transmission must first succeed, and then also be accepted. This structure is hypothesised to reflect the real-world barriers being present at both steps. Each failed share has a chance of rejection and making the receiving agent uncooperative. Uncooperative agents are

labelled ‘resistant’ and their links within the network are disabled. For each agent interaction to be successful, the probability threshold must be exceeded as detailed in section 6.5. Once Initiated, the agent network cycles through sequential quality improvement phases analogous to a plan-do-study-act cycle, with each full cycle proceeding through the pre-set number of iterations.

Each completed iteration incrementally updates progress towards the target improvement state. Stop conditions are set for reaching (i) the improvement target; (ii) a maximum number of iterations, or (iii) falling below a minimum level of network cooperation. Figure 6.4 shows an example of the in-progress reporting through each state from susceptible to adjusted. The x-axis shows elapsed time measured as ‘time units’, an abstract representation of time counting the clock cycles the model has been running. Each ‘tick’ represents an update instruction. The y-axis shows the proportion of nodes in each state updating sequentially across the improvement phases and iterations. In this example, the percentage of resistant nodes (circled) is still well below the maximum level for the experimental run.

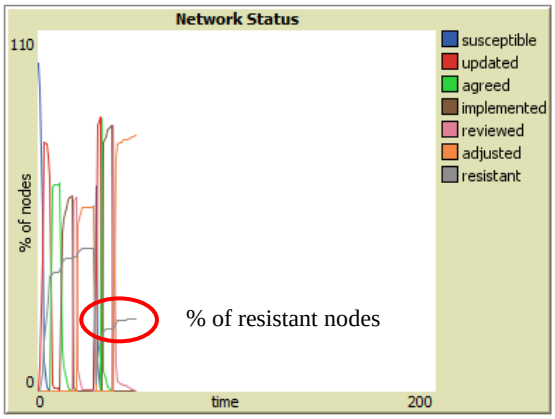


Figure 6.4: Status Reporting Example

Figure 6.5 presents an illustration of the accumulated effect of this process over time. In this example, the impact of the increasing learning gain per iteration is visible in the network status (lower left) and outcome signal (lower right) charts. As time progresses along the x-axis, cycle completion time is faster, and the outcome signal improvement rate is accelerating. Each completed learning cycle improves the chance of exceeding the threshold tests applied to the agent information sharing, meaning there are fewer failed attempts, less opportunity for rejection, and faster progress towards the goal state.

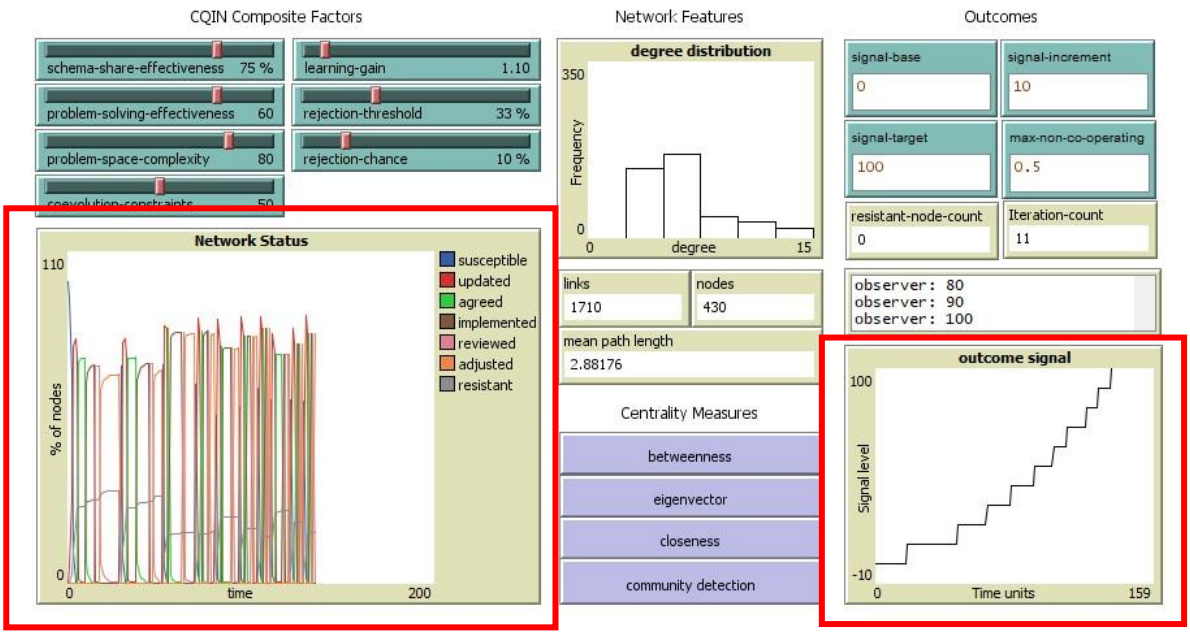


Figure 6.5: CQIN Editing Interface Example 2

Observed at individual agent level, the behaviour is 'share', 'update' or 'reject'. As the model updates over time, the network cycles through successful improvement iterations or eventually encounters a stop condition. Observed at system level, the behaviour is the progress towards the improvement target whilst maintaining a minimum level of agent network cooperation.

As built, the model progresses through the improvement phases manually (update-agree-implement-review-adjust). This separation facilitated code validation and troubleshooting. It also allows individualised weighting to be added to the threshold procedures as a future refinement. For practical use running thousands of replications, the NetLogo 'behaviour space' scripting function runs the iterations as one automated sequence, which is then replicated the desired number of times. Appendix 1 presents the CQIN agent-based model functionality in the form of a unified modelling language (UML) state transition chart.

6.6.3. Network Centrality Measures

In network science there are numerous standardised measures to describe a network structure. Although the majority of these were not used as configurable parameters for this study, they are very relevant when considering a network of adaptive agents engaged in quality improvement.

It is important to re-confirm at this point what exactly is being represented as a ‘network’. Each link between agents is conceptualised as an opportunity to communicate, share ideas, and align understanding and activity. This is an abstraction of the many ways such opportunities actually exist between the diverse professional groups and individuals within the network. The agents themselves are also an abstraction, generic individuals acting in accordance only with their share-update-reject rules. The stochastic nature of the threshold equations is used to mimic the personal and professional heterogeneity of a network of real healthcare practitioners, and the countless opportunities for successful or failed communication. Table 6.2 sets out relevant network measures and their usage within the CQIN ABM.

Table 6.2: Agent Network Attributes Relevant to Information Sharing

Network Attribute	**Description (Ognyanova, 2016)**	**Usage within CQIN**
Node degree	This can refer to *in-degree*; representing the number of incoming connections for the agent; *out-degree*, representing the number of outgoing connections, or *total node degree*, for both incoming and outgoing connections.	The minimum total node degree is configurable
Centrality: Betweenness	A measure of how many times a node features in the shortest path between other nodes.	Able to be measured
Centrality: Closeness	A measure of how closely connected a node is to all the other nodes.	Able to be measured
Centrality: Eigenvector	A measure of how influential a node is, based on the number of connections each of their connections has.	Able to be measured
Community	Groups of densely connected nodes with fewer connections across to other groups.	Able to be measured
Hub	Hubs are nodes with high density of *incoming* connections.	Not used
Authority	Authorities are nodes with high density of *outgoing* connections.	Not used
Homophily	Grouping of nodes based on similar category attributes such as professional group, age, physical location etc.	Not used

For this study, the node degree variable served as an effective high-level parameter to control and vary network density. The *preferential attachment* mode of network construction then created realistic network centrality based on the number of agents and node density. Simply put, new nodes tend to connect to more connected nodes (Barabási,

2016). Given the important role preferential attachment plays in the CQIN behaviour, Barabási's (2016) explanation is quoted here:

Preferential attachment is a probabilistic mechanism: a new node is free to connect to any node in the network whether it is a hub or has a single link. [The preferential attachment equation] implies, however, that if a new node has a choice between a degree-two and a degree-four node, it is twice as likely to connect to the degree-four node (Barabási, 2016, p. 170).

The nature of inter-connection between agents obviously has a critical role in the CQIN model functionality. For the simulation experiments these aspects are controlled via the number of agents, the total node degree, and preferential attachment. However, familiarity with the wider network centrality measures in Table 6.2 becomes extremely valuable as soon as a real-world network is being considered. The lengths of communication paths between agents, the formation of hubs, communities and highly influential nodes is easy to comprehend when describing human agents. Further, the impact that the communication paths has is straightforward to explain. A highly influential node not cooperating will quickly be seen to have a cascading impact on information sharing.

6.7. Verification and Validation

6.7.1. Validation Approach

This section sets out the steps taken to answer the fundamental question – does the CQIN agent-based model function as intended based on the specification? The terms verification and validation can cause some confusion depending on whether the context is epistemological or describing common practice (Oreskes, Shrader-Frechette, & Belitz, 1994). Research involving simulation modelling requires that the epistemological limits are fully acknowledged. For Oreskes et al. (1994), no numerical model can ever be verified in the sense of declared 'true', because this level of verification is only possible in a closed system with all possible parameters known to be correct, such as a mathematical proof (Oreskes et al., 1994, p. 642). However, the function of lines of code can be specified and verified within their defined limits. Similarly, model functionality can be judged valid in the sense of being logically consistent and free from errors, but it does not follow to then claim that the model is a 'valid' representation of real-world phenomena. Making that leap is a logical error of 'affirming the consequent', glossing

over the ultimately incomplete access to real-world information (Oreskes et al., 1994). Windrum et al., (2007) describe the empirical validation of ABM as comparing an unknown, real-world data generation process with a "sufficiently good approximation" of this process within the model (Windrum et al., 2007, p. 5).

Korb, Geard, & Dorin, (2013) approach validation from a different perspective, recommending a Bayesian thought process for ABM validation. This framing emphasises expert opinion as forming the prior probability, achieved in this study through the literature foundations of the conceptual model and iterative feedback from knowledgeable contributors. Empirical data becomes the likelihood, measuring the correspondence between the model and empirical evidence. Prior probability and likelihood combine via Bayes' theorem for a posterior probability.

The model purpose provides another essential input to the testing approach. In contrast to a predictive or decision-making model, CQIN does not attempt to include any case-specific features or a full range of potential inputs. The purpose of testing is not to see how faithfully the model represents the structural details of any given healthcare quality improvement agent network. As a theory-building model, CQIN is a *deliberate* simplification of complex phenomena to facilitate exploration of general theories (Sun et al., 2016). Within these epistemological limits, validation tests serve a legitimate function of assessing a numerical model's outputs relative to the observational data (Oreskes et al., 1994). Practical verification and validation of the model construction is possible and of course necessary.

The specific nature of an agent-based model leads to four specific testing requirements. First, the macro-level system behaviour in CAS arises from the micro-level agent behaviour. *Microvalidation* (of agent interaction mechanisms) is thus essential alongside *Macrovalidation* (of system-level observations) of the model (Wilensky & Rand, 2015). Second, agent-based systems have a path dependence that unfolds in time (Windrum, et al., 2007). Consequently, the initial starting conditions must be established. A third requirement is then establishing what range of system behaviour is being assessed. CAS are sometimes described as non-ergodic, meaning that only a subset of all potential system states are traversed or active (Kauffman, 2000). Assessing *Face Validity* brings these three requirements together. Face Validity asks if the model is a reasonable representation of the real-world system being modelled, based on the specific range of

operational conditions and the observed macro-level behaviour (Wilensky & Rand, 2015; Windrum, et al., 2007). Finally, *Empirical Validation* asks how well the modelled outputs correspond with observed real-world data (Windrum, et al., 2007). This chapter continues by describing the verification of the model code, sensitivity tests, and replication tests. Assessing Face Validity using a real-world example was the objective of the first simulation experiment, and this is detailed in Chapter 7. Empirical Validity of both the CQIN conceptual model and the CQIN ABM is the topic of Chapter 8.

6.7.2. Verification, Sensitivity Tests and Replication

Verification is used here in the meaning of Davies' et al (2007) modelling framework as assessing the reliability of the *computation.* Does the code function correctly? This must test not just the fidelity to a specification or plan, but also whether that specification has captured all the necessary elements. This aspect of model 'completeness' is measured as Content Validity – are the essential factors and relationships from the conceptual model in the computed model? (Korb, Geard, & Dorin, 2013). The careful derivation of the CQIN conceptual model ensured that there was a clear set of requirements (a specification) to operationalise as an ABM.

The recommended practice for building ABMs is to take an iterative approach, working with small subsets of code, ensuring each subset functions as intended by referring back to the expected behaviour of the conceptual model (Wilensky & Rand, 2015). This process was adopted, moving from an initial proof of concept through to more comprehensive and high-fidelity modelling across sixteen full generations of the model. The author is not an experienced programmer; becoming confident in the NetLogo code required construction in discrete modules. Modular construction facilitated *unit testing*, achieved by stepping through each functional module within the code (for example, the initialisation and update functions). The NetLogo visual editing interface allowed mistakes to be identified easily. Constructing the UML representation (Appendix 1) served as a further check of the correspondence between the conceptual model and the functionality as programmed.

Sensitivity analysis tests a model under a range of operating conditions (Wilensky & Rand, 2015). This is both important and difficult for a CAS model. It is important because CAS behaviour is dependent on initial starting conditions. It is difficult not because of the number of variables but because the full range of behaviour was still only

hypothesised at the programming stage. Learning more about the behaviour is the purpose of the model. How is legitimate system behaviour distinguished from programming errors only identified in unusual or transient combinations? NetLogo contains an in-built scripting function allowing repeated operation across any specified range of parameters. Thus sensitivity analysis could be conducted across the range of minimum to maximum values for each variable. The outputs of these test were checked against the hypothesised outcomes of CQIN. For example, lower values of *shared schema effectiveness* should (and did) result in higher agent network resistance. Extreme positive and extreme negative settings produced the respective rapid success and rapid failure outcomes. These were preliminary sensitivity tests; a comprehensive sensitivity test was conducted as one of the simulation experiments presented in Chapter 7.

All tests were replicated sufficient times to assess the natural variation introduced by the stochastic procedure. Replication is important to understand the natural range of outcomes introduced by the randomness in the agent interactions (Wilensky & Rand, 2015). Table 6.3 illustrates this variation for one test. In this test, all the variables remained fixed, and the test was repeated in batch sizes ranging between one hundred replications and two thousand replications.

Table 6.3: Stochastic Variation Across Multiple Repetitions

Number of replications	**Replications achieving full 10% improvement target**	**Percentage achieving target**
100	65/100	65%
200	135/200	67.5%
500	360/500	72%
1000	652/1000	65.2%
2000	1393/2000	69.6%

Table 6.3 reveals the level of variation generated by the intrinsic uncertainty of the model, and the range of this variation, for the given starting conditions. The model outcomes were stable within a range of 65% to 72% at 1000 replications. Extending to 2000 or 5000 replications did not extend the outcome range. Finally, all of the research

experiments were fully reproduced twice, four months apart, with results in the same range of variation.

6.8. Summary of the Agent -based Model

This chapter has described the creation of the CQIN agent-based model and the translation of the underlying theory behind CQIN into a computable simulation model. Clarity about what exactly is being simulated is essential to assessing the model. Given the theory-building purpose, the CQIN agent-based model is intentionally simple and built on the combination of limited interaction rules between agents within a macro-level context of improvement cycles. Interactions between agents are probabilistic, with a stochastic procedure used to create thresholds for successful interactions. The thresholds are moderated by the combined levels of the CQIN variables. Exact identification of the countless reasons for any one cooperation or non-cooperation outcome between agents is not the point. Non-cooperation is premised to occur at some level, and the objective of the modelling is to explore the relative influence of the key constructs and the resulting system behaviour.

6.8.1. Model Availability

The ABM model as developed by the author and used in this study is available to be viewed and downloaded via the NetLogo modelling commons (NetLogo, 2023b). The model, code and research context are fully available for researchers to access, review and extend both the concepts and implementation.

Chapter 7 Simulation Modelling Results

7.1. Introduction

This chapter presents the results of the simulation experiments. Section 7.2 is a description of the experiment conditions and process. Experiment 1, a simulation of the antipsychotic deprescribing project, is the subject of section 7.3. The second experiment is a controlled variation of the *schema share effectiveness* variable, detailed in section 7.4. Experiment 3, a sensitivity analysis, is presented in section 7.5, and a summary and critical assessment conclude the chapter.

7.2. Experiment Protocol

The simulation experiments involved creating networks of cooperating agents, who progress through sequences of improvement (learning) cycles. The system-level output was the resulting progress of each network towards a pre-defined improvement target. Understanding the stop conditions for each experiment is necessary to interpret the results. The sequence stops under two conditions: (i) the programmed number of cycles are completed, or (ii) the number of cooperating agents has fallen below a configurable minimum. If the network stopped prior to reaching the improvement target, this was classified as a failed sequence. Each complete learning cycle was replicated a minimum of 1000 times, ensuring the stochastic variation amongst the iterations converged to the expected range. For this research three separate experimental tests were carried out:

- Test 1: Assess the probability of the face validity case study (deprescribing CAS) achieving its improvement objectives, based on the elicited values from the case study.
- Test 2: Measure the influence of the key *schema share effectiveness* variable.
- Test 3: Conduct a sensitivity analysis using the elicited values from the deprescribing case as the control.

Collectively the three experiments explore the emergent outcome of the agent interactions, the influence of each of the CQIN constructs, and the range of viable values for each of the constructs.

7.3. Experiment 1: Simulation of the Deprescribing Project

For experiment 1, the simulations predicted that the desired 10% reduction in antipsychotic prescribing could be achieved with a probability of 69%, based on the system variables as configured. This percentage is rounded from the direct count of

replications – 688 of the 1000 simulated improvement sequences being successful and 312 failing. A precise predictive value is not the objective, the replication provides an expected range of outcomes as observed in Figure 7.1. Notably, Experiment 1 revealed a substantial non-linear transition zone in the progression towards project success. This transition is shown in Figure 7.1. The y-axis of Figure 7.1 shows progress towards the goal state of a 10% reduction in inappropriate prescribing. The x-axis displays "time units" — an abstract representation of elapsed time using the programme clock cycles. Elapsed time functions as the proxy indicator for the level of learning achieved. Any learning acquired within an improvement cycle improves the starting point for the subsequent cycle, thus problem-solving ability increases over time as more cycles occur.

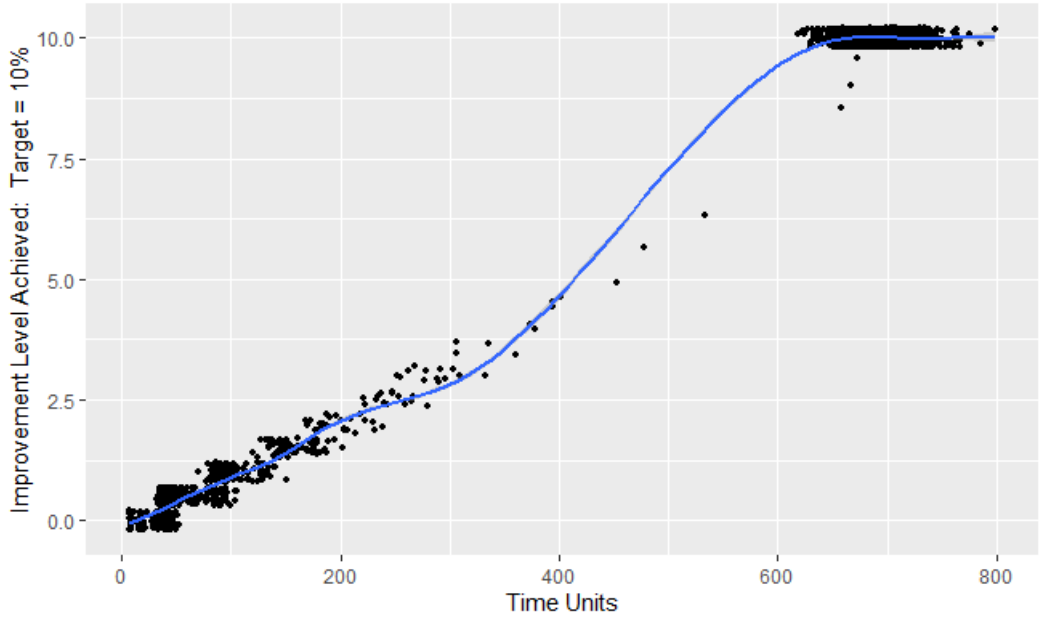


Figure 7.1: Improvement Level Reached by Simulated QI Projects

(Deprescribing project with estimated settings, 1000 simulations)

Each dot in Figure 7.1 is one complete replication of the full learning cycle sequence. The lower left-hand quadrant reveals that those sequences that "fail" predominantly do so very early in the process. As time (and learning) progresses, reaching the target of 10% becomes increasingly likely, and failure becomes rare after about one third of the planned 20 QI cycles have taken place.

7.4. Experiment 2: Impact of Schema Share Effectiveness

Experiment 2 examined the impact of varying the critical *schema share effectiveness* variable across five ordinal levels as shown in Figure 7.2. The levels were set from 10

(very low) to 90 (very high), with 2000 repetitions taken at each level. Meeting the improvement target of 10% is indicated as 'target achieved = TRUE or FALSE'.

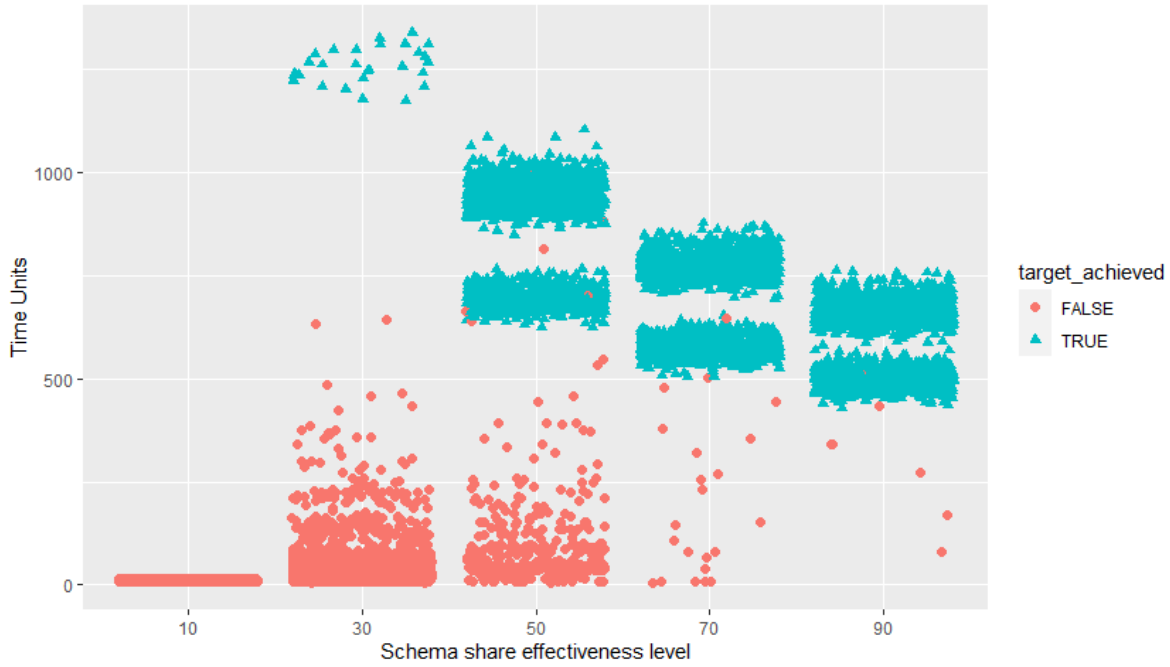


Figure 7.2: Impact of Varying the Levels of Schema Share Effectiveness

Figure 7.2 reveals again the distinct threshold between two system states, with one state being a low probability of success and the second state being a high probability of success. This result indicates that the *schema share effectiveness* variable must meet a minimum threshold for success to be probable, and beyond this threshold failure becomes unlikely. For the given starting conditions, this transition occurs once schema shar effectiveness increases beyond 50%. A second, smaller transition zone is also observed in Figure 7.2 within the successful repetition sub-set. This effect is caused by the time delay introduced when an initial schema share request is unsuccessful, but sufficient agents nevertheless remain active in the network to allow subsequent attempts. These repetitions are therefore ultimately successful but take longer to reach the improvement target.

7.5. Experiment 3: Sensitivity Analysis

Experiment 3 extends experiment 2 to consider the viable range of all the CQIN constructs. Table 7.1 shows the impact of stepwise changes in each CQIN construct. Test configuration 1 is the settings for Experiment 1. The subsequent test configurations alter

one construct in a stepwise manner up and down. The size of the variation was specific to each variable and represented a minimum point of departure from the starting configuration to induce a change. The objective was not at this point to identify the precise level where change occurs, it was to see if change occurred and to what effect. The final row in Table 7.1 shows the percentage of successful repetitions, ranging dramatically from the control value of 69% down to less than 10% and increasing to over 90%. A narrow operating range is revealed for each construct, with limited further value in exhaustive exploration across the full set of possible settings.

The sensitivity analysis reveals that successfully completing a learning sequence is highly dependent on *all* the initial system values. The results shown in Table 7.1 reflect both the non-linear and the holistic nature of CAS. The impact of even small interventions on individual elements is unpredictable, *and* contingent on the settings of other critical elements. Some interventions have non-linear effects (disproportionate to input changes), based on the threshold settings for the variable – for example, varying the *rejection chance* between 5% and 15% moves the likelihood of success from near-certain to impossible (test configurations 6 and 7). These results show that individual change levers can be adjusted, but *all* the variables must be on the right side of their respective thresholds for successful problem solving.

The results for experiment 3 reflect the absence of any smoothing or weighting in the interaction formula. It may be unlikely that a real-world CAS would have such a narrow range of operating viability for each variable. However, an operational 'envelope' for success and joint dependency of the variable is consistent with CAS theory (Holland, 2006, 2012; Kauffman, 2000; Mitchell 2009). One unexpected finding is that two distinct classes of variable are observed. Changes in the levels of *problem-solving effectiveness, problem space complexity, coevolution constraints* and *learning gain* generate success outcomes recognisably proportional to the changes in input level. In contrast, small variations in *schema share effectiveness, agent network node degree* and the *rejection* and *cooperation* thresholds lead to large changes in outcome success.

Table 7.1: Sensitivity Analysis for the Antipsychotic Deprescribing Case

	Test Configurations																
ABM Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
Number of networked agents	410	410	410	410	410	410	410	410	410	410	410	410	410	410	410	410	410
Network type: (preferential attachment 'PA')	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA	PA
Minimum node degree	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Initial signal detection	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42
Signal base	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Signal target	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90	90
Signal increment	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Schema share effectiveness	50%	40%	60%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%
Rejection threshold	33%	33%	33%	25%	40%	33%	33%	33%	33%	33%	33%	33%	33%	33%	33%	33%	33%
Rejection (disengage) chance	10%	10%	10%	10%	10%	5%	15%	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%
Problem-solving capability	50	50	50	50	50	50	50	40	60	50	50	50	50	50	50	50	50
Problem space complexity	85	85	85	85	85	85	85	85	85	80	90	85	85	85	85	85	85
Coevolution constraints	80	80	80	80	80	80	80	80	80	80	80	75	85	80	80	80	80
Learning gain	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	5%	2.5%	7.5%	5%	5%
Planned iterations	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
Required minimum network cooperation	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	50%	33%	66%
Percentage of repetitions meeting full improvement target	**69.0%**	**9.0%**	**94.0%**	**98.3%**	**14.0%**	**99.6%**	**0.0%**	**53.9%**	**75.6%**	**88.3%**	**55.9%**	**71.8%**	**56.2%**	**41.3%**	**79.5%**	**94.6%**	**0.0%**

- Each test configuration consisted of 1000 repetitions.
- Configuration 1 utilised the initial estimated parameters determined in conjunction with project members.
- Configurations 2-17 then varied 1 parameter as a stepwise increase or decrease from the initial value.
- The percentage of repetitions fully meeting the desired improvement objective (10% prescribing reduction) is shown in the final row.

7.6. Summary of the Simulation Experiments

Experiment 1 demonstrated Face Validity of the CQIN Agent-based model, and the sensitivity analysis of Experiment 3 was consistent with the conceptual model and modelling assumptions. All three experiments reflected the anticipated CAS behaviour. The full set of experiments was repeated twice, to ensure that the results could be replicated.

As an exploration of CAS phenomena within a QI context, three important observations with real-world implications for QI were made: (i) nonlinearity; (ii) the importance of the initial starting conditions for all variables, and (iii) macro-level uncertainty arising from micro-level uncertainty. Nonlinearity was exhibited in the progress towards the improvement target and the probability of success. Manipulation of isolated variables certainly has some impact, but all variables must be at viable levels to maintain cooperation. The 'connected web' of variables becomes obvious. The envelope or successful operating range for each variable was surprisingly narrow and it is assumed that the lack of weighting on the variables contributed to this. As noted in the introduction, gathering empirical evidence to weight the variables remains a challenge for researchers.

The uncertainty introduced into the agent interactions, together with the three stop criteria (time, resources, and cooperation level) can progress via agent network cascades to produce uncertainty at macro level output. Even for a set of initial starting conditions that can succeed consistently, the probability of success is not 100%. Clearly, the rejection probability relative to the minimum network cooperation level has the greatest impact on non-linearity. If there is a low probability of dis-engagement, then lower levels of problem solving and shared understanding simply mean that it will take longer to reach an improvement goal.

Chapter 8 Secondary Data Analysis

8.1. Introduction

This chapter presents the detailed qualitative analysis of the 90 selected QI cases. Presenting the analysis findings completes stages IV and V of the Arkey & O'Malley (2005) scoping review framework. The Thematic Analysis is introduced first in section 8.2, covering the coding process. The QCA is detailed in sections 8.3 and 8.4, including the final results. Section 8.5 is a critical analysis of the findings, including support for CQIN and the additional insights from the Thematic Analysis. The chapter concludes with section 8.6, recapping the important findings from the secondary data analysis.

8.2. Thematic Analysis

As introduced in section 5.5, Thematic Analysis is a qualitative analysis method for identifying patterns within narrative data. Reliably identifying themes involves more than simple counts of words or phrases; it requires a close engagement with the subject matter and constant assessment of the text meaning in relation to the theme (Joffe, 2011; Vaismoradi & Snelgrove, 2019). The defined CQIN constructs functioned as the thematic criteria for the analysis, given the primary objective of examining empirical evidence supporting CQIN. Each of the cases was read in full, then re-read and annotated for thematic coding. This coding was then entered into a tracking spreadsheet and reviewed against the original article. Analysis and coding took place over 18 months, and consistency checks were made between cases from different stages of the process. The content and meaning had to be examined in context. An abstract concept such as *schema share effectiveness* is not directly labelled as such in QI reports. However, the underlying theme of *building shared understanding of goals and actions amongst participants* can be identified, as the following examples illustrate (case identification (ID) in parentheses). The use of a case ID number is adopted to efficiently reference the cases; a mapping table between each case ID and the full bibliographic reference is provided as Appendix 3.

CQIN Construct: Schema share effectiveness

- *The aim is for everyone to understand their role in a child's care plan.* [41]
- *Our use of PDSA cycles demonstrated the significance of working collaboratively across teams. Building productive working relationships between teams required*

recognising strengths and weaknesses and working together to reconfigure services in a way that was patient centred. [43]

- *A one-day summit was held involving approximately 40 key stakeholders including physicians, house officers, nursing staff, and pharmacists, as well as representatives from transportation services, environmental services and bed flow management*. [46]

CQIN Construct: System signals

- *Email communication on comprehensive health assessment rates was sent to case managers each week* [42]
- *In summary, data and feedback processes were described by some participants as a means to initiate and implement change* [82]
- *The introduction of a simple, real-time bedside reporting tool and risk management process facilitated active engagement with families and patients, and the opportunity for disclosure and learning for staff.* [63]

The risk of confirmation bias, or simply finding what was hoped for, was minimised by adopting a conservative approach to coding, recording only clear and unambiguous examples. Because coding was limited to one researcher, there was limited opportunity to seek further interpretation of ambiguous or very subtle references to themes. The trade-off with this approach is that some examples will certainly have been missed. In that sense, the thematic coding can be viewed as 'low resolution'. Latent references were certainly captured but only where they were clear and obvious. Appendix 4 contains coded examples for each of the CQIN constructs.

Joffe (2011) makes the point that there is little value investing in the effort of thematic analysis working only with pre-existing themes, and not making use of the rich qualitative data each case provides. Seven themes not originally included in CQIN, that refine the model or are otherwise relevant to the research question, were identified. These additional themes are discussed in detail in section 8.5 (the QCA results are presented first to provide full context to the findings).

8.3. Qualitative Comparative Analysis

Once all cases were coded a preliminary data matrix was prepared setting out the presence or absence of the CQIN constructs within each case. The subset shown in Figure 8.1

illustrates this, with each row describing a case and the presence or absence of the construct. The full data matrix is included as Appendix 5. Each row contains the set of conditions for improvement success, as interpreted from the thematic analysis. They can be thought of as low resolution 'recipes' for success.

Case ID	Schema share effectiveness	Agent network	Problem solving effectiveness	System signals	Autonomy	Learning gain	Coevolution constraints	Focus / Time	Problem space	Option generation	Interpersonal factors	Budget	QCA 1/0	Count
26	1	1	1	1	1	1	1	1	1	0	0	1	1	10
7	1	1	1	1	1	0	1	1	0	1	1	0	1	9
14	0	1	1	1	1	1	1	1	1	1	0	0	1	9
8	1	1	1	1	1	1	0	1	0	0	1	0	1	8
16	1	1	1	0	1	1	1	1	0	0	0	1	1	8
20	1	1	1	0	1	1	1	1	0	1	0	0	1	8
24	1	1	1	1	1	0	0	1	0	1	1	0	1	8
25	0	1	1	1	1	1	1	1	0	1	0	0	1	8
70	1	1	1	1	0	1	1	0	1	1	0	0	1	8
3	1	1	1	0	0	1	0	1	1	1	0	0	1	7
4	1	1	1	0	1	0	0	1	0	0	1	1	1	7
38	1	1	1	1	1	0	0	0	1	1	0	0	1	7
41	1	1	1	1	1	1	0	0	0	0	1	0	1	7

Figure 8.1: Sample of Cases Showing Presence (1) or Absence (0) of CQIN Constructs

The constructs are arranged left to right by frequency of observation. 72% (65/90) of the cases have 4 or more CQIN constructs present. Only 1 case was recorded without some combination of the first 3 constructs — *schema share effectiveness*, *agent network* or *problem solving effectiveness.*

A need to provide more granularity within the CQIN constructs became apparent at the commencement of the thematic analysis. In practice, descriptions of the CQIN constructs often referenced a subset or incomplete range of constituent parts covered by the high-level construct. As an example, *problem solving effectiveness* can be described in various ways, and the theme of *autonomy* was apparent in the cases. Extra granularity allowed for better matching of the language and concepts reported in the cases with subsets of each high-level construct. To ensure consistency, the recommended approach for constructing a BN was adopted to provide this granularity, expanding out the original CQIN constructs using an appropriate reasoning idiom (Fenton & Neil, 2019). In this instance, the *definitional* idiom was applied, asking what are the important things that make up *schema share effectiveness* or *problem solving effectiveness?* The objective at was not to be exhaustive, but to identify the main contributory factors to the higher-level construct. In a BN context, these contributory factors become the parent nodes of the core

nodes. An example of this relationship is shown in Figure 8.2. The contributory factors of *budget*, *focus/time* and *autonomy* can also be synthesised into a combined concept - *energy/resources.* Ultimately the synthetic nodes were not required, but they are described here to illustrate how the BN nodes can be aggregated or deconstructed. The result is a very flexible map of connections, where the granularity can be adjusted to match the research objective and manage the complexity of the BN (Fenton & Neil, 2019).

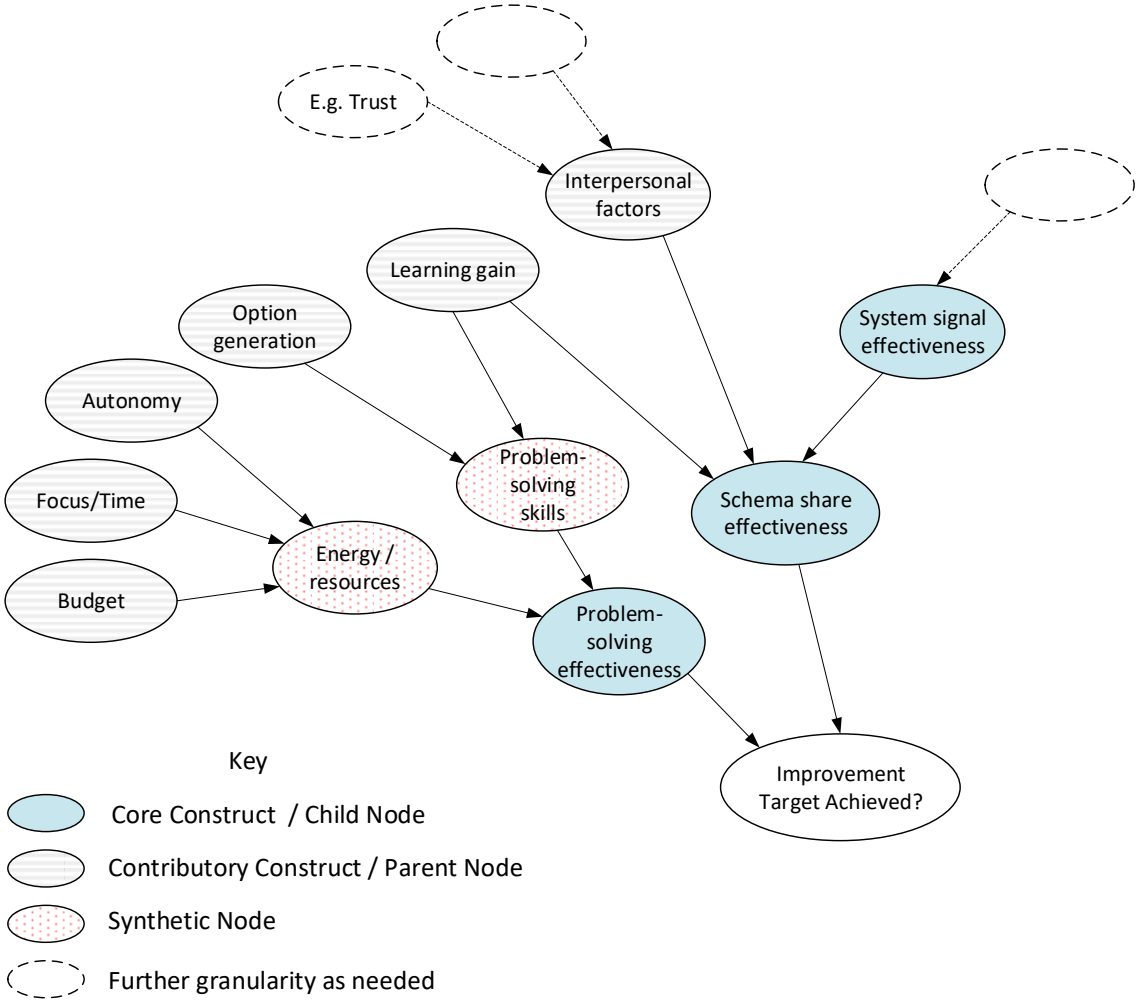


Figure 8.2: Example of Core and Contributory Factors

For the thematic analysis, it quickly emerged that reported QI interventions often used the granularity level below the core CQIN constructs. For example, interpersonal factors such as trust and collegiality, or problem-solving factors such as autonomy. Table 8.1 lists each of the core and contributory CQIN constructs, together with the observed coverage level of these across the 90 cases.

Table 8.1: Presence of CQIN Constructs Within Cases

Functional Level	**CQIN Construct**	**Proportion of Cases Including Construct**
Level 1: Core	Schema share effectiveness	0.69
Level 1: Core	Agent network	0.66
Level 1: Core	Problem solving effectiveness	0.63
Level 1: Core	System signals	0.51
Level 1: Core	Coevolution constraints	0.32
Level 1: Core	Problem space	0.34
Level 2: Contributory	Autonomy	0.33
Level 2: Contributory	Learning gain	0.34
Level 2: Contributory	Focus / Time	0.28
Level 2: Contributory	Option Generation	0.27
Level 2: Contributory	Interpersonal Factors	0.24
Level 2: Contributory	Budget	0.14

Table 8.1 reveals strong occurrence of the *schema share effectiveness, agent network,* and *problem solving effectiveness* constructs, in approximately two thirds of cases. *System signals* was identified as a condition in 51% of the cases. There is then a gap down to 34% or less for the remaining constructs. For *problem space* and *coevolution constraints*, it is possible that the inhibiting effect of these concepts is less likely to be noted prominently in typical QI reports. Ideally, a high complexity context and coevolution constraints are absent rather than present. Overall, the published QI reports tended to focus on what positive steps were taken for success, and hence focus on the problem solving actions, often without in-depth consideration of how these steps directly overcame the underlying barriers such as coevolution constraints. Equally, it is possible that these constructs were simply not strongly present in many cases. The contexts may

not have been genuinely complex in the sense that this research is examining. It is nevertheless encouraging that these two constructs were still able to be consistently recognised in approximately one third of cases.

The contributory constructs are all represented in a range between 24% to 33% of cases, except for budget, which was simply not often discussed or described directly. As noted, the conservative approach taken to identifying themes may be underrepresenting the true frequency of latent constructs. The proposed relationships between the core and contributory constructs were examined and supported by a paired analysis, as seen in Figure 8.3:

	Schema share effectiveness	Agent network	Problem solving effectiveness	System signals	Autonomy	Learning gain	Coevolution constraints	Focus / Time	Problem space	Option generation	Interpersonal factors	Budget
Schema share effectiveness		45	38	31	20	22	22	20	20	17	21	10
Agent network	45		34	29	25	21	23	22	18	16	19	11
Problem solving effectiveness	38	34		29	23	23	18	19	23	17	13	9
System signals	31	29	29		18	17	13	11	16	16	7	4
Autonomy	20	25	23	18		12	10	14	12	13	9	8
Learning gain	22	21	23	17	12		12	9	11	11	8	4
Coevolution constraints	22	23	18	13	10	12		10	8	7	8	7
Focus / Time	20	22	19	11	14	9	10		9	7	8	7
Problem space	20	18	23	16	12	11	8	9		10	1	4
Option generation	17	16	17	16	13	11	7	7	10		4	1
Interpersonal factors	21	19	13	7	9	8	8	8	1	4		4
Budget	10	11	9	4	8	4	7	7	4	1	4	

Figure 8.3: Paired Analysis of CQIN Constructs

Within Figure 8.3 each row-column intersection is the count for both these constructs being present within a case. The highest association for *interpersonal factors* was with the *schema share effectiveness* and *agent network* constructs, consistent with the posited

parent-child relationship between these constructs. Similarly, *learning gain* is strongly associated with *problem solving effectiveness* and *schema share effectiveness*. Figure 8.3 also provides a clue as to the interpretation of the *problem space* and *coevolution constraint* variables. *Problem space* is most strongly associated with *problem solving*, suggesting that they may have been described as two sides of the same concept. Similarly, *coevolution constraints* were most strongly associated with the *agent network*. For those cases reporting attention being paid to the agent network communication paths and building shared schema effectiveness, *coevolution constraints* may well have been overcome without being explicitly identified.

8.4. Truth Table

The data matrix was converted to a truth table using Tosmana 1.61 QCA software (Cronqvist, 2019). Specifically the analysis performed was *crisp set* QCA (csQCA, referred as QCA in this study). Crisp set QCA is Boolean, where the conditions of interest are true or false. Fuzzy set QCA and multivariate QCA were considered but found to be unsuitable for this data. The fundamental barrier was inability to define any credible criteria for a ranked scale consistently across such a diverse range of QI reports. Even a simple high/medium/low distinction would introduce a subjectivity that could not be supported by the data when reviewed at such a distance from the actual activity. For consistency of interpretation, if the secondary data cases reported improvement progress towards a goal, with some substantiation in reported qualitative or quantitative data, quality improvement had occurred (i.e. the improvement outcome was logically 'True').

At this point an important aspect of the QCA process is addressed. Conventionally, QCA requires a good balance of positive and negative outcomes (Rihoux, & De Meur, 2009). Without negative outcomes, determining *necessary* conditions is not possible, as there are none (or very few) cases to prove that without the given condition, the outcome is negative. A *necessary* condition *must* be present for the outcome to be true. Due to the strong publication bias effect, the secondary data was very unbalanced, with only 3 cases reporting negative outcomes. For this reason no attempt is made to identify necessary conditions for QI success using this analysis. However, multiple positive cases still allow for consideration of *sufficient* conditions. If the outcome is true or virtually always true when a configuration of conditions is present, then that configuration is sufficient. This approach is consistent with one of the originating motivations behind QCA (Ragin, 1999; Rihoux & De Meur, 2009). The recommendation is for qualitative researchers to shift

their focus from searching always for *necessary* causal conditions, and to examine combinations of causal conditions to explore *sufficiency* in more depth (Ragin, 1999). This idea is well suited for complexity research, where each case can legitimately be considered unique.

The QCA software constructs a truth table, and then performs a Boolean minimisation. This process was performed multiple times to determine which conditions (CQIN constructs) to include. The combination of 'low resolution' reporting and unbalanced data set (87/90 with positive outcome) resulted in an excessively large truth table. Restricting the conditions to the four positive core CQIN constructs resulted in a more manageable table, justified by the parent-child reasoning between the constructs. For example, if *problem solving effectiveness* was present, it is reasonable to assume that the contributing factors *option generation* or *autonomy* were also present. The final version of the truth table was constructed using the four most frequently observed CQIN constructs: *schema share effectiveness, agent network, problem solving effectiveness*, and *system signals.* Table 8.2 is the resulting truth table. Ragin (1999) provides the distinction between the truth table and the raw data matrix. Rather than simply recording presence or absence of the CQIN constructs across cases, the truth table presents 15 qualitatively distinct causal combinations (Ragin, 1999).

Table 8.2: Truth Table of Cases and Core CQIN Constructs

	Case ID	Schema share effectiveness	Agent network	Problem solving effectiveness	System signals	Outcome	Coverage
1	C39, C48, C49	0	0	0	0	0	0.03
2	C28	0	0	0	1	1	0.01
3	C52, C55, C58, C63, C66, C90	0	0	1	0	1	0.07
4	C21, C34, C61, C72	0	0	1	1	1	0.04
5	C36, C60	0	1	0	0	1	0.02
6	C6, C53, C82	0	1	0	1	1	0.03
7	C44, C47	0	1	1	0	1	0.02
8	C10, C11, C12, C14, C25, C33, C56	0	1	1	1	1	0.08
9	C1, C17, C42, C46	1	0	0	1	1	0.04
10	C18, C29, C30, C43, C79	1	0	1	0	1	0.06
11	C5, C37, C62, C64, C77, C78, C81, C86	1	0	1	1	1	0.09
12	C13, C23, C27, C54, C59, C69, C74, C80, C83, C84, C89	1	1	0	0	1	0.12
13	C2, C15, C19, C50, C68, C71, C73, C76, C88	1	1	0	1	1	0.10
14	C3, C4, C9, C16, C20, C22, C31, C32, C35, C45, C57, C65, C67, C85, C87	1	1	1	0	1	0.17
15	C7, C8, C24, C26, C38, C40, C41, C51, C70, C75	1	1	1	1	1	0.11

Each row in Table 8.2 shows the secondary data cases where the combination of causal conditions was True, along with the improvement outcome as True or False. The

last column is the coverage, a measure of the empirical relevance. Coverage reports the degree to which a necessary or sufficient condition explains instances of the occurrence of the outcome (Rubinson, 2019). For example, the combination of *agent network* and *schema share effectiveness* being True, with *problem solving effectiveness* and s*ystem signals* being False, covered 12% of the cases (Row 12).

The 15 causal configurations are made up of 4 prime implicants, which are the minimum distinct configurations associated with the outcome, derived from pairwise comparison and reduction (Ragin, 1999). Given that each construct included is an implicant, this does not add any useful information. However, the implicants do minimise one step further to provide one complete configurational solution for the secondary data as analysed:

Solution:
Agent network{1} + Problem solving effectiveness{1} + System signals{1} → Quality improvement

The solution states that "If the *agent network* condition is True, OR the *problem solving effectiveness* condition is True, OR the *system signals* condition is True, the outcome of *quality improvement* will be True. This solution reveals that *schema share effectiveness* was never present on its own; it is always present in combination with one of the other CQIN constructs.

8.4.1. Consistency and Coverage measures

Consistency describes the relationship between each causal configuration and the outcome (Rubinson, 2019). The absence of contradictory cases (multiple outcomes for the same causal combination) indicates consistency, subject to the acknowledged caveat regarding the unbalanced data set. *Sufficiency* is reasonable to assume given the number of cases reporting positive outcomes and lack of contradictory outcomes, but *necessary* causal configurations cannot be identified using this data.

Coverage, a measure of empirical validity, provides evidence for the CQIN constructs, reporting the occurrence of each causal configuration within the empirical data. The QCA analysis is summarised visually in Figure 8.4. Each rectangle in Figure 8.4 represents one combination of conditions from the truth table. In the top left rectangle

are the three cases where none of the CQIN constructs were identified, indicated by the string [0000]. The bottom right rectangle lists the cases where both schema share effectiveness and agent network were identified, shown as [1100]. Overall, Figure 8.4 displays that the 90 secondary data cases can be represented via 15 causal configurations.

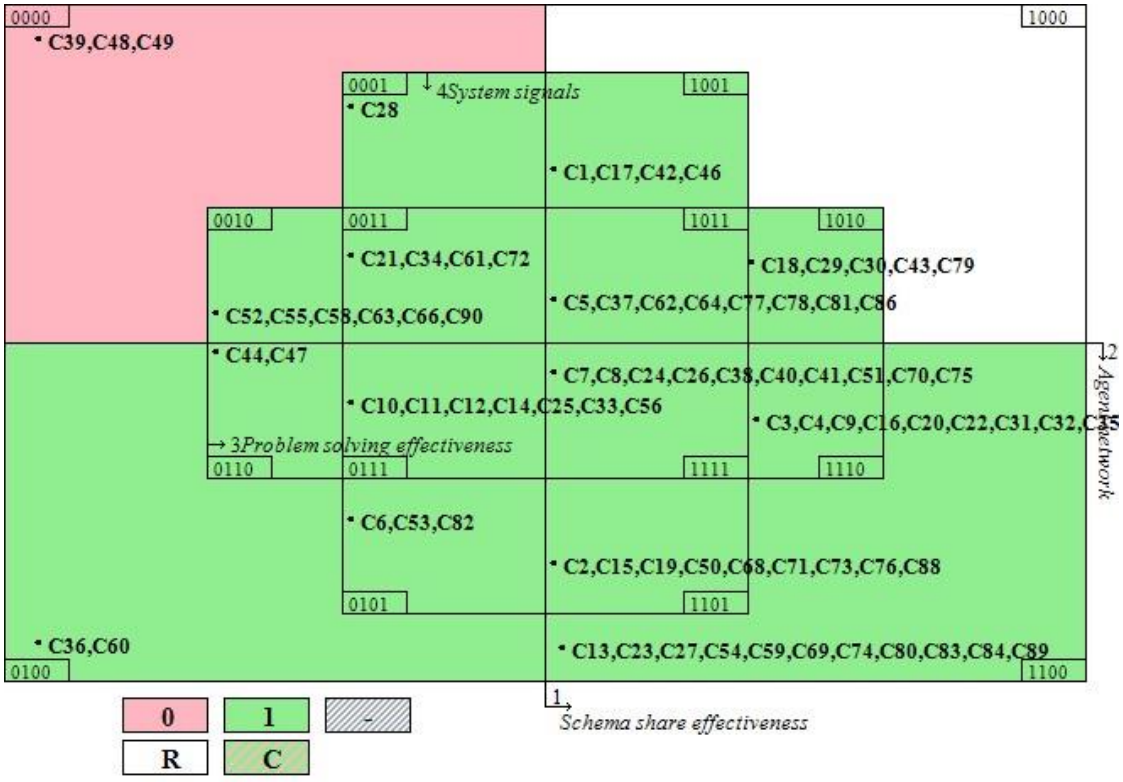


Figure 8.4: QCA Venn Diagram of Cases and Core CQIN Constructs

8.5. Critical Analysis

In research involving multiple case studies, the question of 'how many cases are sufficient?' should be addressed. Small (2009) recommends seeking *saturation*, where case replications should continue until a point is reached where nothing new is being added to the theory and analysis from the data being collected. The secondary data *n* of 90 case is considered an intermediate size *n* in the context of QCA (Berg-Schlosser, et al., 2009). Saturation was reached relatively early, with minimal variation in results when compared at 30, 60 and 90 cases.

The 3 negative cases (cases 38, 48, and 49) were rare instances of researchers reporting (and succeeding in getting published) QI activity that was unsuccessful. The Boolean nature of the QCA process merged the absence of a construct, not describing a

construct, and inadequacy or problems with a construct. Each of these three scenarios would indicate the construct as 'False' for that case. However, the narratives were clear on the underlying reasons for failure, suggesting poor implementation (low problem solving effectiveness), high complexity (an inhibiting and inadequately negotiated problem space), lack of alignment (low schema share effectiveness) and under-resourcing (lack of focus/time and budget).

8.5.1. Support for CQIN

The secondary data analysis provides empirical support for the validity of the CQIN constructs and the underlying conceptual model. All the constructs were identified, even with the conservative approach to latent thematic description. Three of the six core constructs were identified in two/thirds of the cases, and a fourth, *system signals*, was identified in half the cases. The posited relationships between core and contributary constructs was supported by the paired analysis and absence of contradictory outcomes.

At least four cases (Cases 3, 84, 88 and 89) described *systematic* frameworks to interpret QI and complexity, analogous to CQIN. These examples show that QI practitioners and researchers are starting to move beyond selective references to complexity elements and constructing more coherent frameworks and models for QI specifically in the context of complexity.

8.5.2. Additional insights

In addition to empirical support for CQIN, the Thematic Analysis identified 7 new themes as relevant considerations for CQIN and the wider research question: (i) applying or extending QI tools within a complex context; (ii) signal timing/integration; (iii) coordination/navigation; (iv) professional skills and capability development; (v) complexity-informed leadership; (vi) workflow integration, and (vii) self-directed learning. This section examines these themes in more detail.

(i) Applying QI tools within a complex context

(Cases 18, 25, 26, 28, 41, 43, 46, 47, 51, 55, 65, 72)

The suitability of QI methods and tools within complex contexts was a topic of discussion noted early in the literature review. Are such methods completely unsuitable or can they be informed by or adapted to complex contexts? This theme suggests that practitioners are applying a wide range of conventional QI approaches

as intentional responses to complex scenarios. The example below refers to a team using rapid PDSA cycles to increase their learning and build a shared schema:

> *Through discussion and reflection on practice within the project team, we developed a driver diagram to show how we envisaged service configuration.* (Case 43, p.3).

(ii) Signal timing/integration

(Cases 1, 5, 6, 19, 33, 47, 61, 62, 76, 81)

An additional and valuable dimension to the system signals construct is the element of *timing*. When are the signals available to respond to and how smoothly do they integrate into existing workflows?

> *Without an accurate view of demands and risks to quality, staff cannot respond to problems in a timely or effective way.* (Case 5, p. 207).

> *Participants consistently identified weekly feedback of audit results, allowing prescribers to suggest interventions and provision of the laminated reminder card, as the most important components.* (Case 33, p. 4).

(iii) Coordination/navigation

(Cases 15, 23, 27, 36, 40)

Several cases focussed strongly on the communication and handovers amongst multiple teams of clinicians and support staff. At an extreme, this involved the creation of dedicated roles to navigate the complexity of care pathways spanning teams, facilities, and providers. The implication goes beyond supporting the development of robust communication paths between very large agent networks; it suggests a legitimate problem-solving response is to create roles that can learn and navigate system complexity.

> *The challenge of coordination is increasingly interpreted as rooted in complexity of care systems (Braithwaite et al., 2017). This view suggests that care is routinely organised through large numbers of heterogenous groups cooperating in non-linear patterns of interaction, rather than through relatively well defined, linear pathways.* (Case 15, p.1).

(iv) Professional skills and capability development

(Cases 16, 17, 32, 37, 38, 77, 83)

This theme adds another valuable dimension not included in CQIN. Cases emphasising this theme put forward the idea that focusing on professional development of all kinds — clinical, organisational management, people leadership and quality improvement skills — increases knowledge and therefore improves problem solving ability.

> *The teams learn the tools and methods of QI and each team selects a project that aims to solve a problem in its own clinical setting….* (Case 16, p. 252).

(v) Complexity-informed leadership

(Cases 3, 29, 30, 38, 73, 87)

This theme provides empirical support for the complexity frameworks emphasising the role of leaders in navigating complexity such as *Cynefin*, *Adaptive Leadership*, or *Complexity Leadership*. Each of these frameworks highlight the role of leadership and the need for leadership *actions* to be appropriately matched to the challenges posed by complexity (Kurtz & Snowden, 2003; Heifetz, et al., 2009; Uhl-Bien & Marion, 2009).

> *A leadership style that embraces complexity science, accepting that barriers are part of change, and that perseverance and patience are necessary attributes. Adopting new ideas or readiness to alter flow requires buy-in from the point-of-service team members.* (Case 38, p. 284).

(vi) Workflow integration

(Cases 20, 52)

This theme expands on the signal timing/integration theme to encompass the QI interventions. The finding is that tight coupling between interventions and operational workflow increases the chances of the intervention working.

> *The use of a departmental huddle translates into situational awareness.* (Case 20, p. 827).

(vii) Self-directed learning

(Cases 8, 14, 24, 29, 31, 32, 43, 83)

This final theme extends the professional development theme, suggesting that learning is also facilitated by having the opportunity to *apply* new skills.

> *Moreover, we modified the LSS method to iteratively add new information to the process maps as we engaged new stakeholders until we seemingly achieved some level of saturation - until no new information appeared.* (Case 31, p. 72).

8.6. Summary of the Secondary Data Analysis

The Thematic Analysis and QCA revealed strong empirical support for the CQIN model. The CQIN constructs were observed in the empirical data, despite the conservative approach to coding. The core concepts of *agent network communication paths, schema share effectiveness, problem solving effectiveness* and *system signal integration* were all strongly represented. The posited relationships between the core constructs, and between each core construct and associated contributory variables, were also supported by the lack of contradictory cases. The secondary data thus support Content and Face Validity for CQIN . The close engagement with the secondary data, specifically the published reports of QI initiatives by practitioners seeking improvements under complex conditions, added valuable detail to the CQIN model. One key finding, consistent with literature reviews, is that the references to complexity are still of very uneven quality (Thompson et al., 2016; Turner & Baker, 2019). A second notable finding was the widespread application/extension of conventional QI methods. The actions of practitioners provide an alternative viewpoint to the argument that these methods are *intrinsically* unable to cope with complexity.

The thematic analysis revealed insights beyond evidence in support of CQIN. Several of these suggested or added detail to the constructs, such as the timing and workflow integration of system signals (as opposed to the mere presence of these signals). In contrast, complexity-informed leadership was an unavoidable theme within the cases, and this leadership is not obvious anywhere in the CQIN model, and within CAS more generally. This important finding is discussed in depth in Chapter 11.

The QCA demonstrated that 90 QI cases studies, from 22 countries and encompassing 41 healthcare focus areas, could be represented as 15 causal configuration

sets. Questions arise when considering the degree of generalisability for any ‘successful’ complex QI initiative. On the one hand, if each case is unique, it is unlikely that simply copying a solution from elsewhere will work. But the CQIN levels suggest a category distinction between the high-level features of CAS and any specific interventions. The secondary data analysis strongly suggests that CAS-aware ‘strategies’ are consistent across a wide range of successful healthcare QI.

Chapter 9 Bayesian Network Analysis

9.1. Introduction

Chapters 7 and 8 presented the simulation modelling and secondary data analysis results. This chapter completes the analysis phase by integrating the conceptual model, agent-based simulation findings, and the secondary data findings into a Bayesian network (BN). Representing CQIN as a BN allows the relationships between model constructs to be examined using a causal structure whilst making use of the simulation and empirical data to support the variable relationships and probability assumptions. A brief description of BN is provided first in section 9.2. Section 9.3 then outlines the development process for the CQIN BN structure presented in this work and section 9.4 describes the node probability tables. Section 9.5 details calibration of the model, incorporating expert judgement to account for the unbalanced data set used to train the model. A comprehensive range of modelled results are presented in section 9.6, including verification and internal validation. A consideration of CQIN now represented as a Bayesian network is provided in section 9.7. A critical assessment, section 9.8, concludes the chapter.

9.2. Bayesian Networks

Bayesian networks, also known as *belief* networks, are a method to formally represent conditional probability and uncertain relationships amongst a set of variables (Puga, Krzywinski, & Altman, 2015). Using Bayes' theorem, Bayesian probabilistic reasoning describes the process of updating our prior belief about an uncertain hypothesis in light of new evidence. Our initial belief is termed prior probability, while our updated belief is termed posterior probability (Fenton & Neil, 2019).

A BN consists of a directed acyclic graph (DAG), which describes the connections between each node (variable) in the network, and an associated node probability table (NPT) for each node within the graph. The nodes represent each variable and the edges between nodes connect dependent variables. The connections are directed, meaning that there is a direction of influence between nodes. An edge from node A to node B represents a causal or influential relationship that node A has on node B. Node A is often described as a parent of node B. Acyclic means that loops are not permitted in BN, because no node can be parent of itself. The NPT is the probability distribution of the node, given the

parent nodes. For nodes without parents, the NPT is the probability distribution of the node based on frequency of occurrence (Fenton & Neil, 2019).

The connections between nodes are built from three fundamental relationships as shown in Figure 9.1:

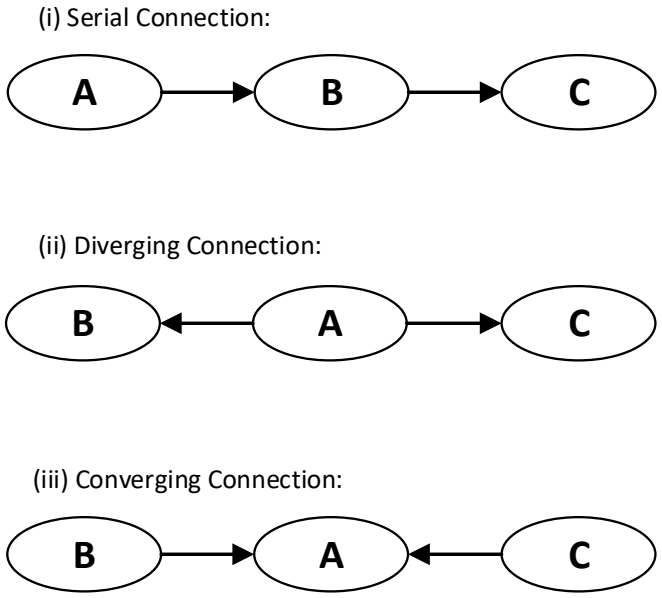


Figure 9.1: Fundamental Node Connection Types

With serial or causal connections there is a sequence of connections in one direction between the nodes, and a causal assumption from A to B and then to C. With a common cause or diverging connection, node A is the common cause of node B and node C. The third connection type is the common effect or converging connection, where multiple parent nodes converge on a child node. In this configuration, nodes B and C converge on node A.

From these three elementary node connection types, the hypothesised or observed relationships between nodes can be built up to form the full DAG. An important feature of the connections is that they also show which nodes are not connected and therefore not directly dependent. This allows efficient calculation of the conditional probabilities and provides a visual check of the causal assumptions (Fenton & Neil, 2019).

9.3. Constructing the DAG

Recommended practice for constructing BNs is to build up network fragments using appropriate reasoning idioms. The use of idioms has been shown to result in consistent

representation of the underlying relationships between nodes given the reasoning archetype (Neil, Fenton & Neilson, 2000; Fenton & Neil, 2019). General reasoning idioms include notions such as *cause-consequence*, *definition*, and *measurement*. The idea of reasoning idioms has also been applied to medical and clinical decision-making. Medical reasoning idioms allow the interaction or influence of clinical concepts such as *pathology, comorbidities, risks* and *complications* to be modelled in a consistent manner (Kyrimi at al., 2020). To build the CQIN Graph, the starting point was the hypothesised relationships between the core CQIN constructs as introduced in Chapter 4 (Figure 4.1). Following the ABM experiments and the secondary data analysis, the original relationships were modified very slightly to incorporate changes suggested by the data. A direct causal path between the *system signals* and the *improvement outcome* nodes was added, given this direct relationship was suggested by the secondary data analysis. An indirect path between *schema share effectiveness* and *improvement outcome*, via *problem solving effectiveness*, was also added. This combination of direct and indirect causal paths for this variable better reflected the conditional configurations observed in the comparative analysis.

One final adjustment was required to make the translation from the agent-based modelling context, where the agent network was literally coded into the model, to a representation of the agent network performance within the BN. The critical interpretation was narrowed to one question: *Is the communication across the agent network enabling or inhibiting?* This Boolean node allowed for a generalisable interpretation of the many different network engagements and activities as described in the secondary data cases. This translation into a True or False outcome also made the *network viability* outcome variable from the agent-based modelling redundant. The final graph for the BN CQIN model is shown in Figure 9.2. The labels A to I identify the CQIN constructs as originally presented in the conceptual model. The node fragments are constructed based on causal reasoning, for example, *problem space complexity* influences the level of *improvement achieved*. The definition reasoning idiom is also applied, where parent nodes are identified. An example of this idiom is where *interpersonal factors* contribute to *schema share effectiveness.* The CQIN BN graph reflects the intended parsimony of the conceptual model, including only those factors initially posited or subsequently refined during the research process. Nevertheless, the model structure supports the addition of further levels of detail if required for theoretical or practical applications.

9.4. Building the Node Probability Tables

Early versions of the CQIN BN treated the variables as ranked nodes, using a five-point ordinal scale ranging from 'very low' to 'very high'. Although intuitive and feasible, this approach was subsequently difficult to support using either expert judgement or empirical data. The experience working with the antipsychotic deprescribing project team confirmed the difficulty of capturing assessments reliably and consistently across multiple experts. What constitutes 'very low' rather than merely 'low'? This finding is consistent with the literature on soliciting expert judgement in decision making generally, and particularly when interpreting probabilities (Fenton & Neil, 2019; Morgan, 2014). The interpretation challenges are reduced if the node probability distributions are reconceived as Boolean (i.e. True or False). 'Was this condition present or absent?' The potential for bias is not eliminated and some form of real-world data is still required to validate the NPT values. The secondary case study data provided an opportunity to do this, and the Boolean nature of the crisp-set QCA data further strengthened the decision to use Boolean probability tables for the CQIN BN. The Boolean approach is also consistent with the interaction thresholds used within the agent-based modelling, leading to Boolean outcomes – was the threshold achieved or not?

One important consequence of using the secondary case data within the BN is the treatment of non-observed or unknown values. Each variable in the BN has three potential states – True, False or unobserved. In the data, there are only two states – True (reported as being present), and False (not reported). This results in a configuration decision for the NPT expression. An unobserved value can be assumed to be false (as it was not reported), or it can be treated as genuinely unknown. Using a logical AND function for the NPT expression assumes each input is False unless confirmed as True, and all inputs must be True for a True output. Using a logical OR function reduces the impact of non-observed inputs, better reflecting the range of solution 'sets' within the secondary data cases.

The raw data matrix from the secondary data analysis, once relabelled with the BN node labels, was successfully imported into AgenaRisk (www.agena.ai) to populate the NPTs. However, the very unbalanced dataset and the conflation of 'not present' with 'not observed' conditions limited the usefulness of the model trained on data only. A second model was developed, calibrating the empirical data with expert judgement.

9.5. Process to Incorporate Expert Judgement

The starting point for calibration was the model version trained exclusively on case data. This version was problematic for two reasons. First, selection/publishing bias resulting in unbalanced data with only 3 of 90 cases self-reported as "unsuccessful". Second, the variation across the different solution 'sets' within the cases, with factors present in some sets but absent in others. This problem is made worse when combined with the 'low resolution' of observations (i.e. did Case X simply not report anything about 'system signals', for example, which may have been present and contributing but not reported or missed in coding).

An important success criterion for any refinement was to retain the real value in the information from successful cases – acknowledging the reported presence of the variable and the relative proportions of influence across many different cases. Combinations of variables supports the hypothesized parent-child relationships in BN. Retaining the idea of different solution 'sets' also reflects the empirically observed variation across the QI cases. Additional success criteria include minimum manual intervention, consistency, and transparency of the calibration steps.

Three calibration steps were implemented:

1. A *noisy-OR* function was applied to incorporate expert judgement that retained the relative weighting of inputs, and allowed for different input combinations to influence the outcome of the node

In the noisy-OR function each parent variable has a weighting value between 0 and 1. The weighting is the probability that the outcome will be true if this parent variable, and no other variable, is true (Fenton & Neil, 2019). A 'leakage' value is also included, which is the probability that the outcome state of the node will be True when *none* of the parent variables are True. Noisy-OR functions have been shown to effectively model many combinations of parent variables (Fenton, Noguchi, & Neil, 2019).

The data-trained node probability tables on the *schema share effectiveness* and *problem solving effectiveness* nodes were updated with a noisy-OR function, retaining the weightings within the data:

Schema share effectiveness

Noisy-OR (interpersonal factors, 0.24, learning gain, 0.34, network communication paths, 0.33, 0.05)

Problem solving effectiveness

Noisy-OR (autonomy, 0.34, schema share effectiveness, 0.66, budget, 0.14, Focus_Time, 0.28, option generation, 0.27, learning gain, 0.34, 0.05)

The BN nodes without parents retained the frequency of True/False observations as learned from the imported data.

2. The *Agent network communication paths* weighting on *Target improvement level achieved node* was reduced.

The *agent network communication paths node* was overly influential in the first model. This was because it has no parent nodes, a direct path to the outcome, and high prior distribution based on relative frequency within the cases. Using only the relative frequency as the weighting is too high as it is based on published cases only. The outcome node weighting for this input was reduced from 0.69 to 0.33

3. Final Outcome Node - Target improvement level achieved?

For the final outcome node, the input weighting for the four positive core constructs was reduced to compensate for the publication bias. The objective was to set weighting levels that would better represent a prior probability of 0.5, or essentially neutral in terms of probability of success or failure. At the same time, the reworked weights must still allow for all the findings from the secondary data to be recreated. For the positive constructs, this reduction was approximately half the original value. The two negative constructs, *problem space complexity* and *coevolution constraints*, were also reduced but not so substantially. This is because each of these constructs was under reported when compared to the others.

From:

Noisy-OR (problem-solving effectiveness, 0.63, schema share effectiveness, 0.69, agent network communication paths, 0.66, coevolution constraints, 0.32, problem space complexity, 0.34, system signals,0.51, 0.05)

To:

Noisy-OR (problem solving effectiveness, 0.33, schema share effectiveness, 0.33, agent network communication paths, 0.33, coevolution constraints, 0.2, problem space complexity, 0.2, system signals, 0.25, 0.05)

Calibration Results

These three changes reduced the prior probability of success to a more realistic level, from 90% for Model v20 (fully derived from the training data) down to 53% for model v21 (just slightly better than even money if nothing else is known). The other scenarios within model v21 now behave as expected. The final combined NPT, with this mix of learned and knowledge nodes, was able to replicate the secondary cases and hypothesised scenarios without problematic or confounding results.

The QCA results contained multiple solution sets and missing/missed observations. Incorporating these data into the BN raised an intriguing question. Do *all* the parent variables need to be True for the outcome variable to be True? If so, an AND function is required. If the output can be True provided at least one of the parents is True, an OR function is implied. The QCA results may be an artifact of the QCA coding process, but they do suggest an OR function within the real-world cases. This means that provided one of the QCA implicants is True, for example, *system signals*, then the improvement outcome will be True in at least some of the cases. One interpretation is to assume that all other necessary conditions are also True but simply not observed. Alternatively, the notion of improvement might be specific to the unique problem space being navigated. For some scenarios, non-observed variables may genuinely not be influential. Guided by the QCA analysis, the CQIN BN was configured as a Noisy-OR function; it is a simple configuration to reset this to use an AND function instead. The impact of this decision is shown in the examples below, with the OR function biasing the model slightly towards a high marginal probability of success (i.e. the prior probability, with no observations made).

9.6. BN Model Results

9.6.1. Verification

Verification in the form of conformance to specifications, as discussed in Chapter 6, is supported by traceability of the model constructs back to the conceptual model and the agent-based model. Sensitivity analysis was conducted by using extreme values to test that the model functioned as expected.

The initial or prior probability for *target improvement level achieved*, with no additional information observed, is shown in Figure 9.3. Next, Figure 9.4 shows the CQIN BN model with all inputs set to True. Figure 9.5 presents the inverse, with all inputs set to False. As expected, with all inputs True, there is virtual certainty the outcome is True. The two extreme scenarios are not exactly symmetrical, due to the unbalanced data set (biased towards success) and the OR function variation. Nevertheless, with all input conditions False, there is an 88% probability the outcome will also be False.

Another preliminary test of the BN behaviour can be made by setting a value for the effect or outcome node and checking that the resulting back propagation updates the probability distributions of the parent nodes as expected. This is shown in Figure 9.6, where the *target improvement level achieved* node is set to False, and the respective values for the parent nodes reflect the expected weightings from the QCA.

9.6.2. Validation

Face and content validity for the CQIN BN are supported by two independent steps. First, the development process via the conceptual model and the ABM ensured essential features based on the theorised model were included. Second, the ability to import case data and represent the empirical cases demonstrates that the node relationships and probability tables are supported. Concurrent validity, an assessment of whether the model is consistent with other independent tests, is also supported by the congruence between the ABM and BN results (McLachlan et al., 2024).

Validation using empirical data was conducted by setting observations from the secondary data cases. Case 32 was a typical example, and the reported set of conditions for Case 32 are shown in Figure 9.7. Note that where a parent and child node are both set with an observation, the child node observation blocks the parent observation, which is redundant (Fenton & Neil, 2019). In Case 32, the autonomy observation is redundant due

to the *problem solving effectiveness* observation being set. The reported success of Case 32 is shown in the high probability of success for *target improvement level achieved.* Case 49 (Figure 9.8) shows the example of negative observations as identified in this case, with the corresponding low probability of success at 21.55%.

Figures 9.9 and 9.10 show the positive and negative extremes for just the four positive constructs exactly as used in the QCA analysis. With all four nodes negative, the model predicts an 82.75% probability of failure. The positive example predicts an 81.33% probability of success. These examples are BN representations of the QCA truth table introduced previously in Table 8.2. Tables 9.1 and 9.2 (Section 9.7) add the BN predictions to the QCA results.

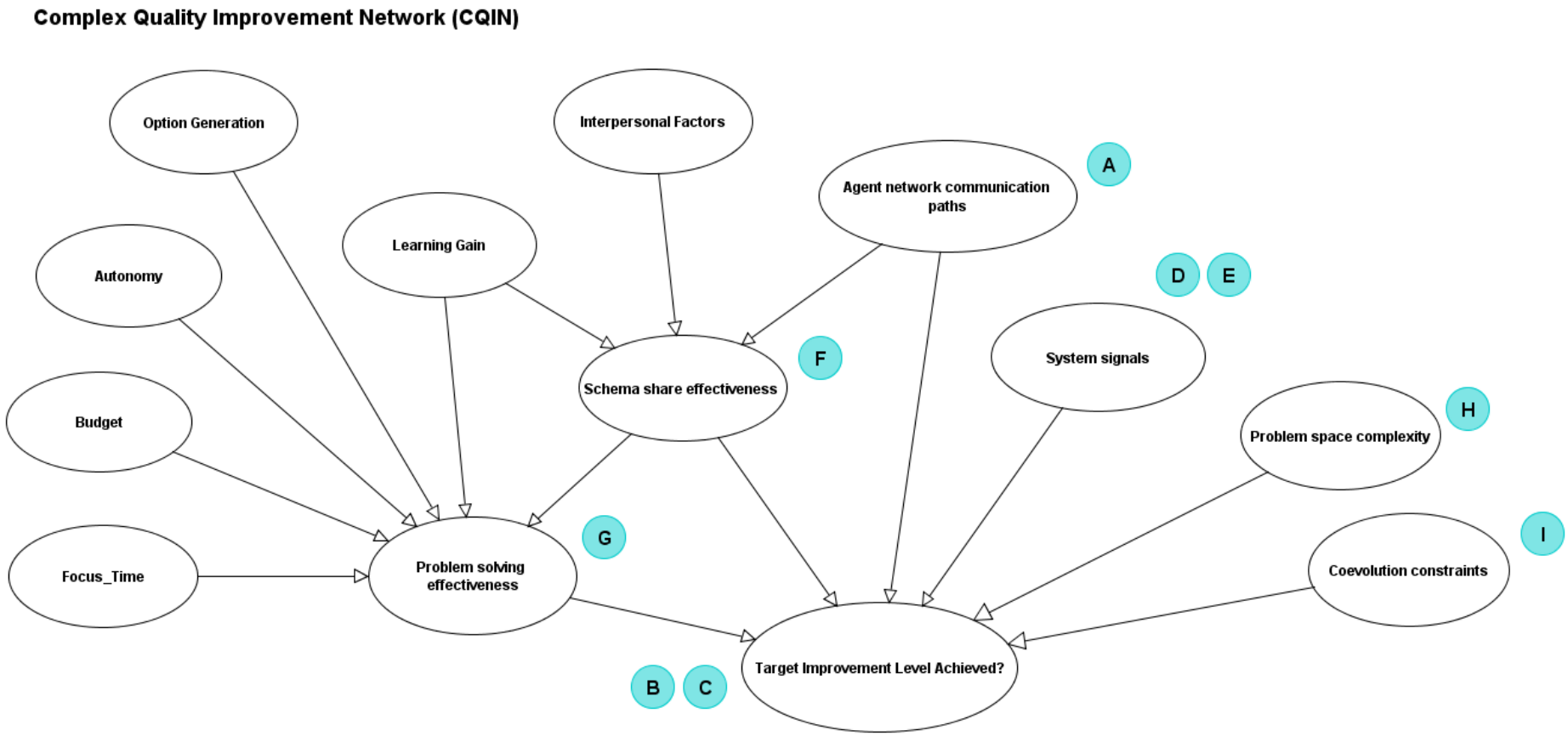


Figure 9.2: Directed Graph for CQIN Bayesian Network

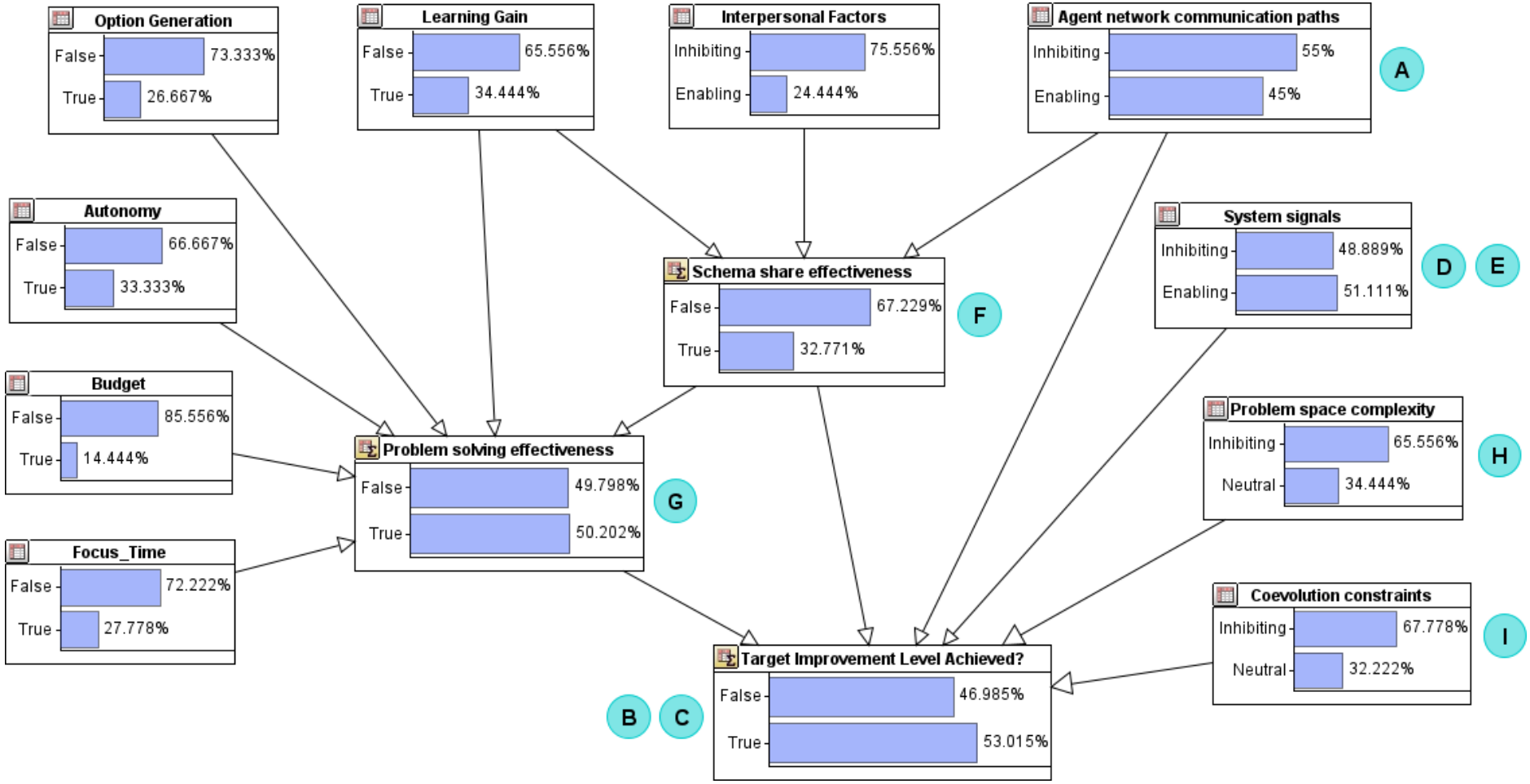


Figure 9.3: Prior Probability with No Observations Entered

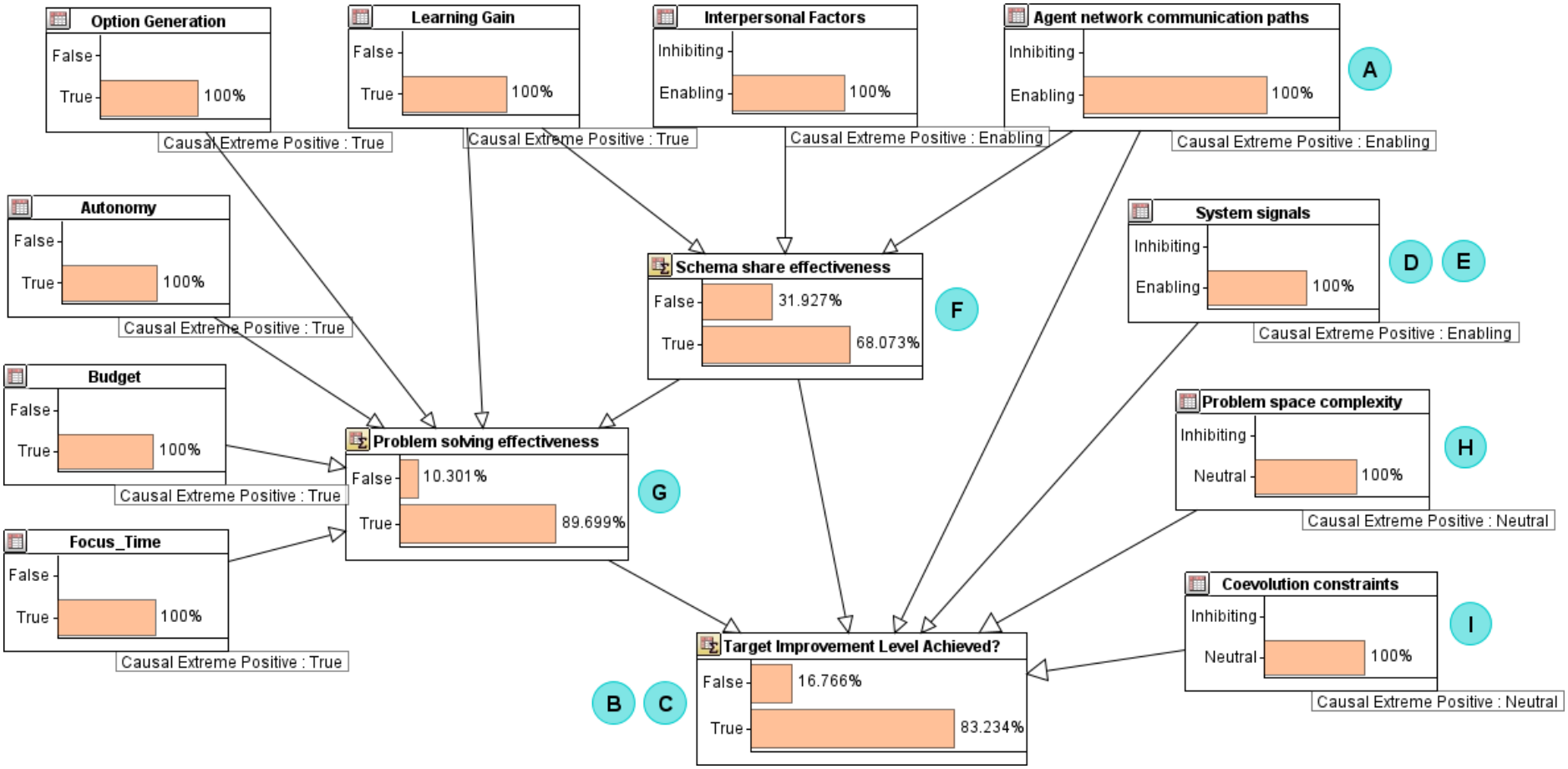


Figure 9.4: Extreme Positive Settings for CQIN BN

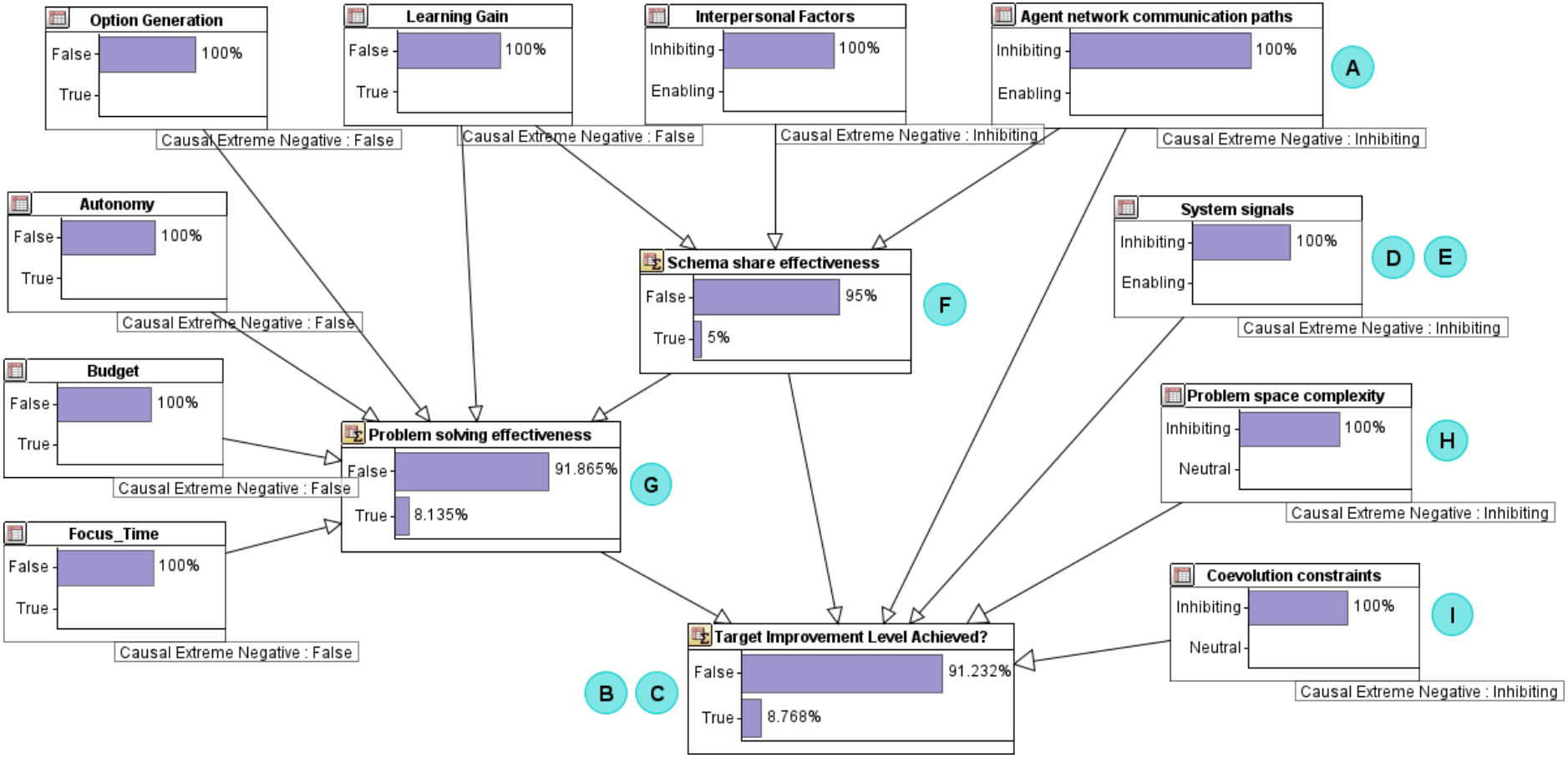


Figure 9.5: Extreme Negative Settings for CQIN BN

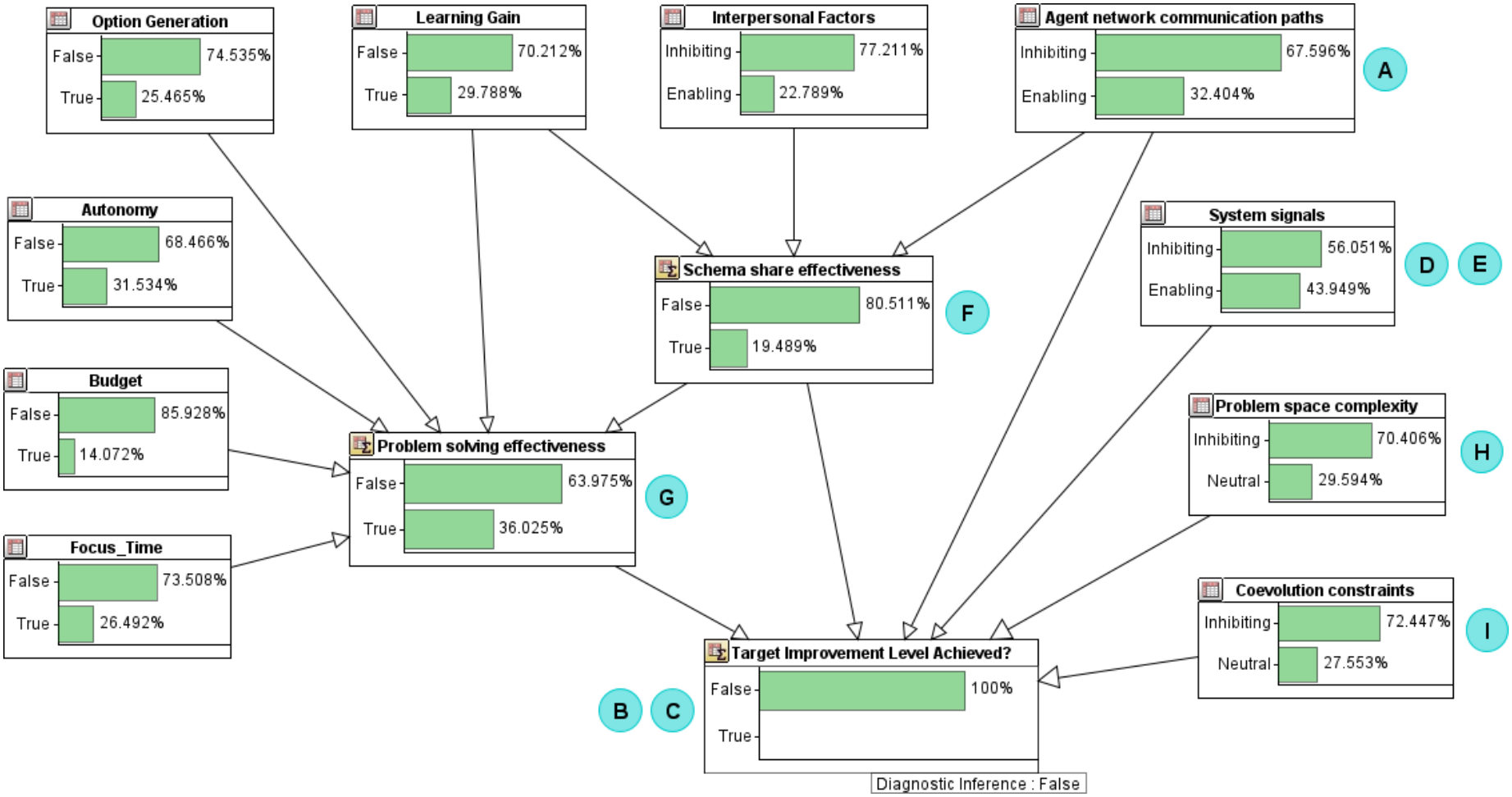


Figure 9.6: Diagnostic Inference Example for CQIN BN

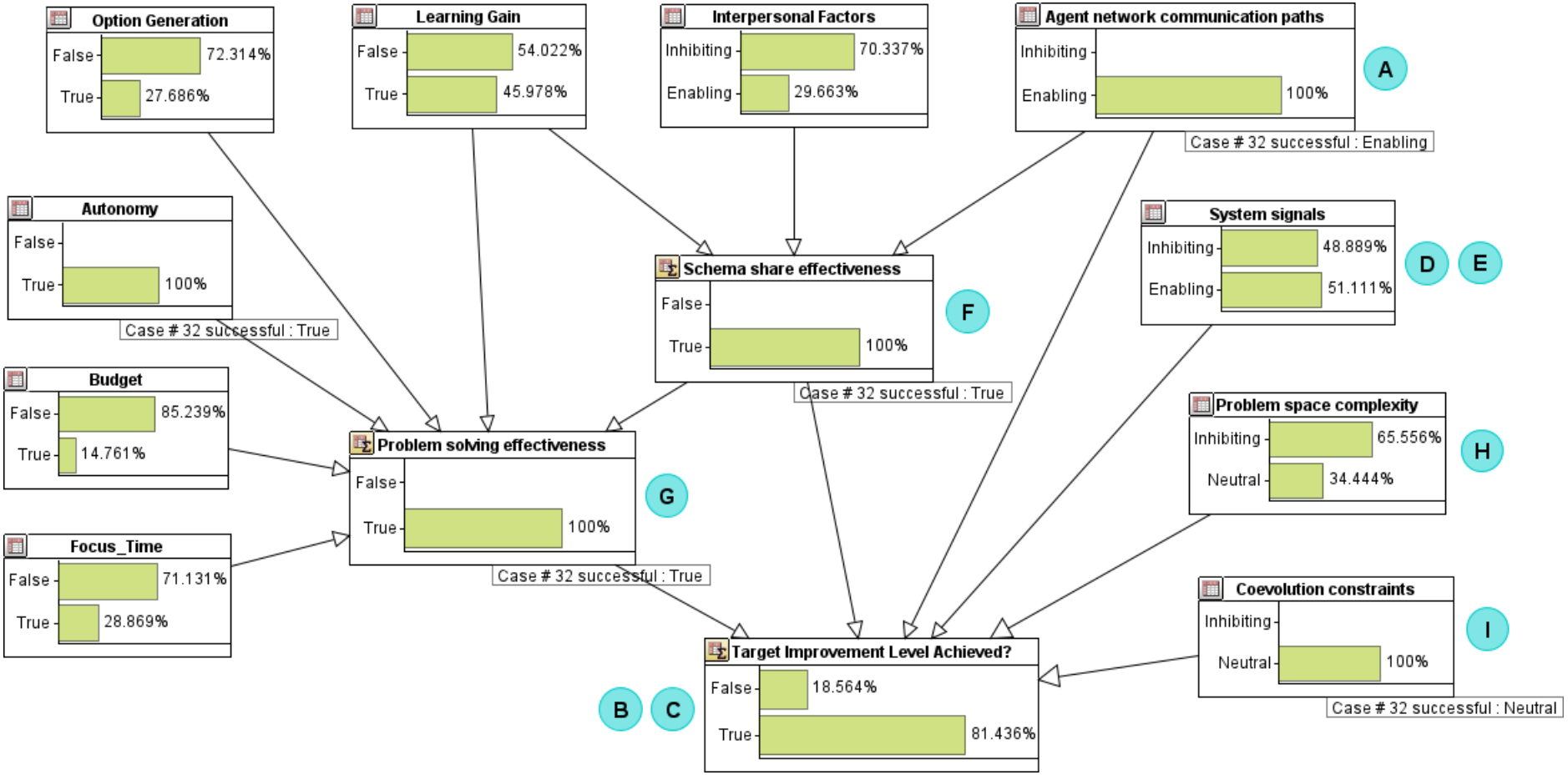


Figure 9.7: Secondary Data Example Case #32

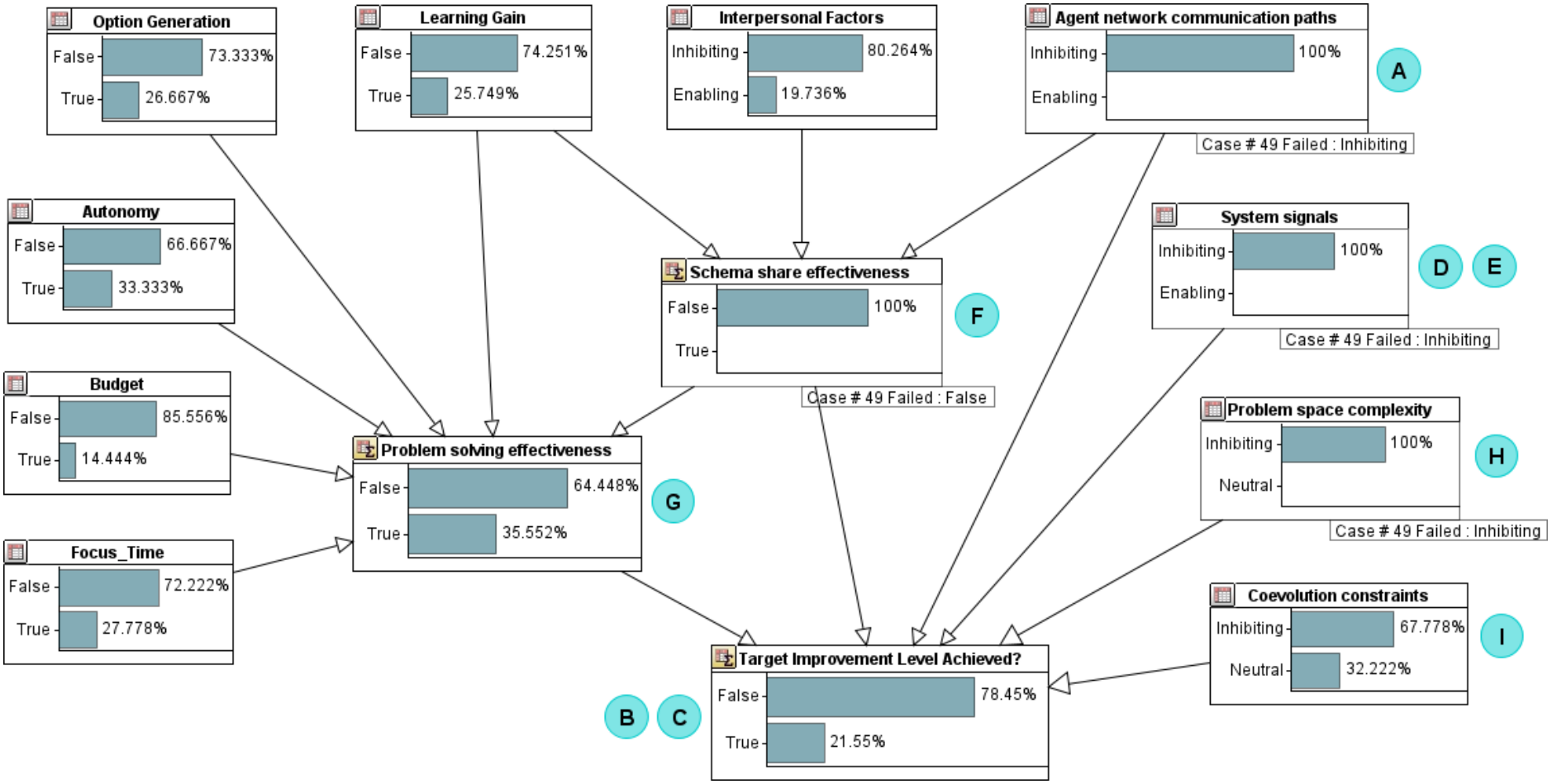


Figure 9.8: Secondary Data Example Case #49

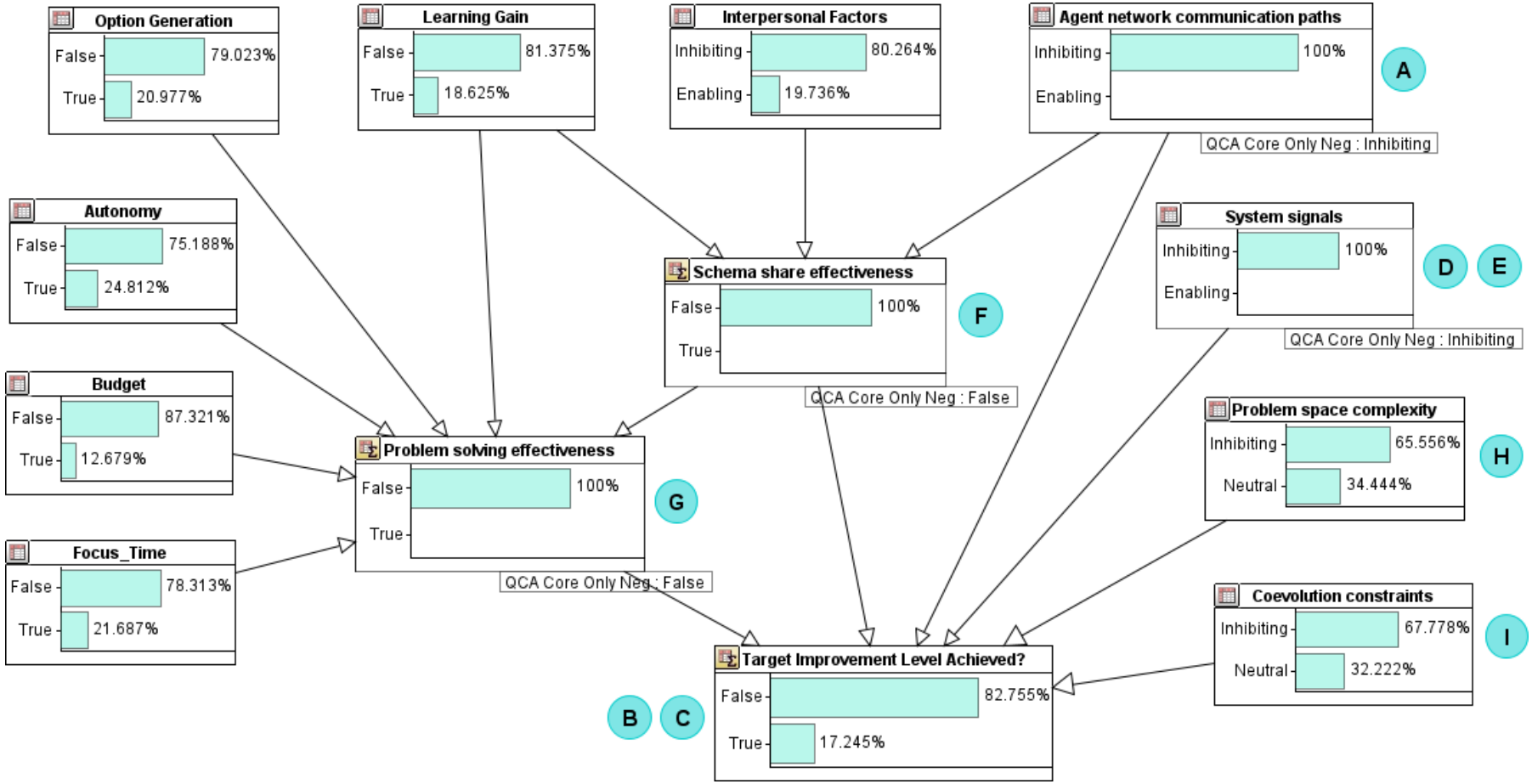


Figure 9.9: Four Core CQIN Constructs Observed as False

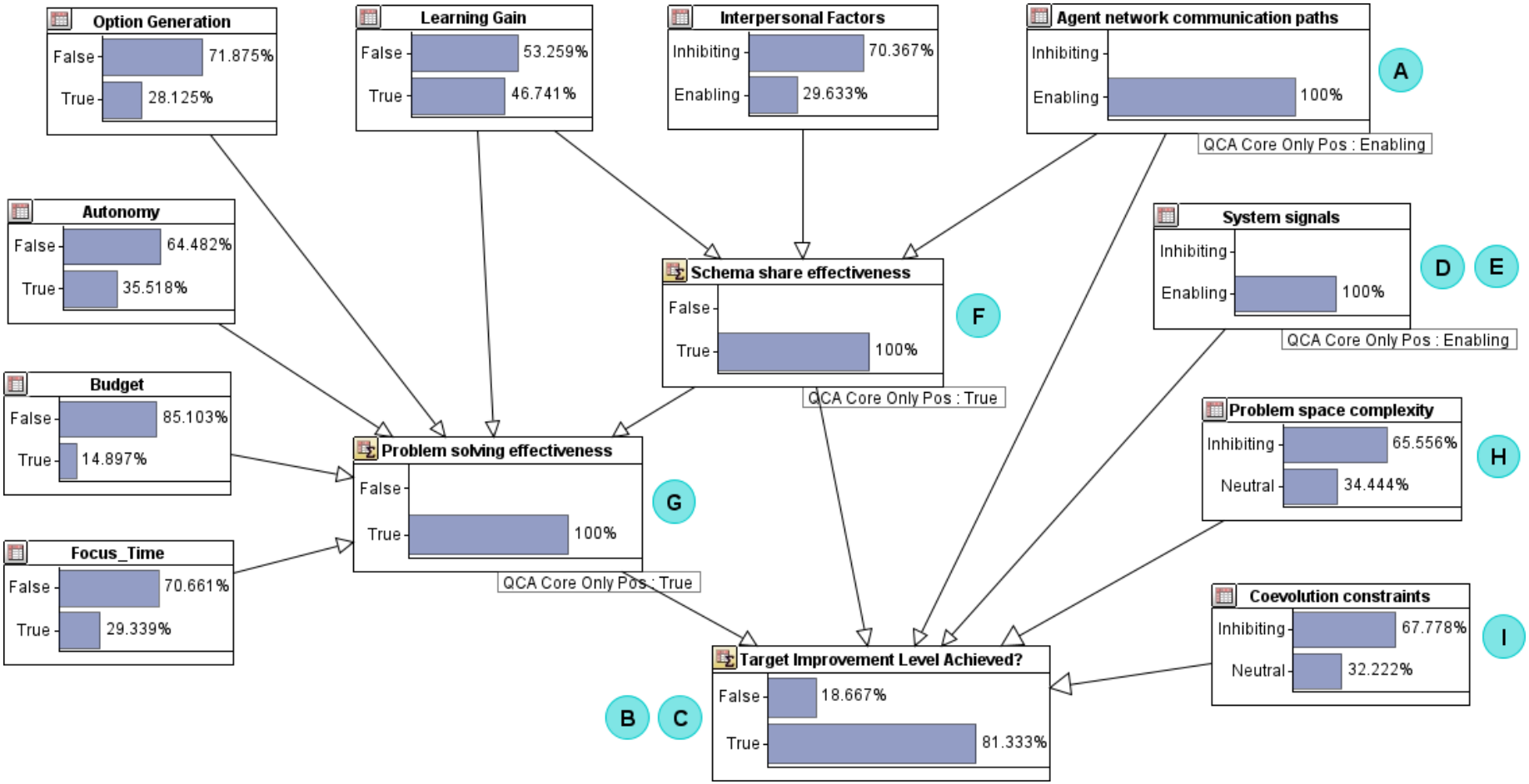


Figure 9.10: Four Core CQIN Constructs Observed as True

9.7. CQIN Represented as a Bayesian Network

Figures 19 to 27 show the CQIN BN in a range of observed states. The CQIN BN brings together the originally hypothesised variable relationships and the results from the secondary data analysis to model QI as CAS behaviour. The relative influence of each CQIN construct is shown in Figure 9.11 as a Tornado chart:

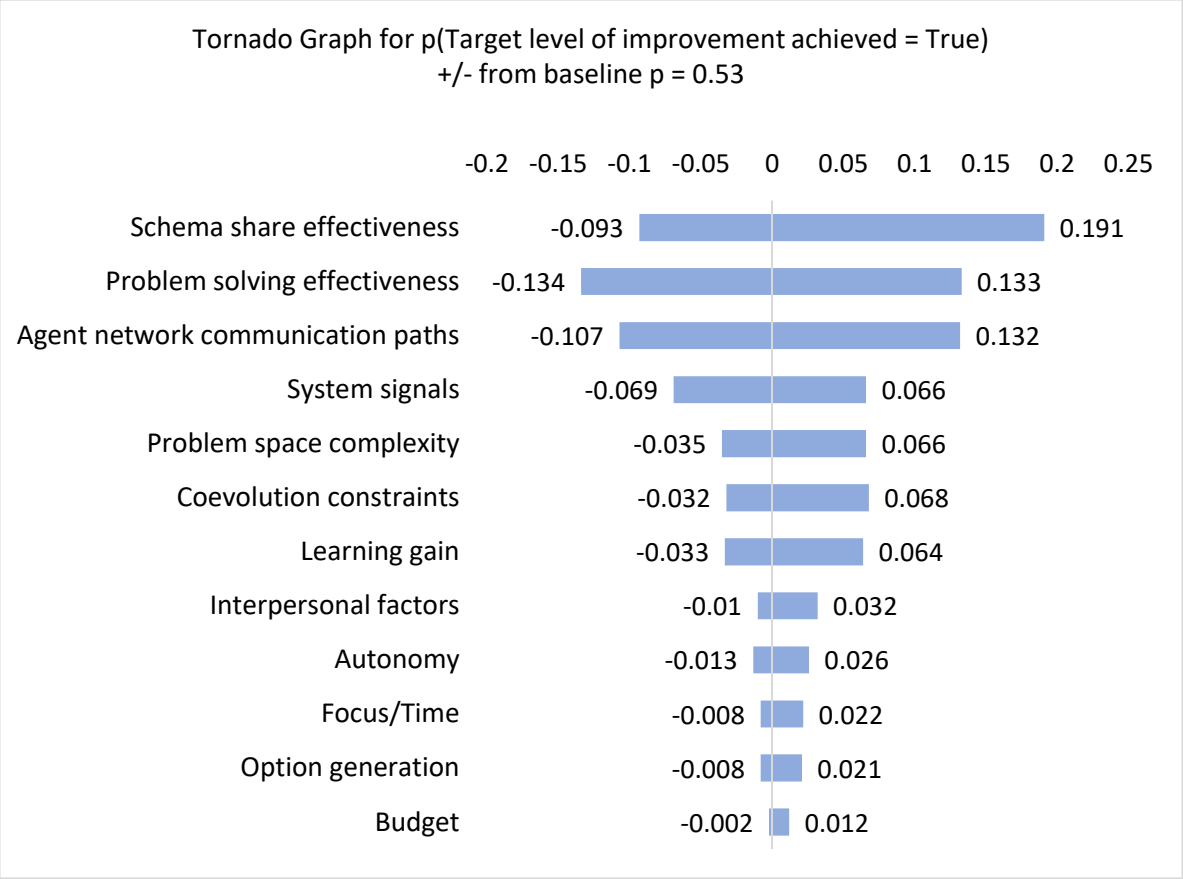


Figure 9.11: Tornado Graph of CQIN Constructs

Figure 9.11 shows the relative influence of each CQIN construct, for example the *schema share effectiveness* construct has a 9.3% negative effect on the marginal probability of a successful outcome if False, and a 19.1 % positive effect if True. Similarly, the *agent network communication paths* construct has a 10.7 % influence if False and a 13.2% influence if True. The relative influence of the core (child) and contributory (parent) CQIN constructs is conveyed. Combinations of values for each construct increase the conditional probability based on the specific combination, such as the case examples provided in Figures 9.7 and 9.8.

The overall representation of the secondary case data within the CQIN BN is now shown in Table 9.1. The BN predictions are consistent with each combination of the four core CQIN constructs.

Table 9.1: QCA Truth Table with BN Prediction of Success

	Case ID	**Schema share effectiveness**	**Agent network**	**Problem solving effectiveness**	**System signals**	**Outcome**	**Coverage**	**BN Prediction of Success**
1	C39, C48, C49	0	0	0	0	0	0.03	0.17
2	C28	0	0	0	1	1	0.01	0.59
3	C52, C55, C58, C63, C66, C90	0	0	1	0	1	0.07	0.66
4	C21, C34, C61, C72	0	0	1	1	1	0.04	0.74
5	C36, C60	0	1	0	0	1	0.02	0.66
6	C6, C53, C82	0	1	0	1	1	0.03	0.71
7	C44, C47	0	1	1	0	1	0.02	0.75
8	C10, C11, C12, C14, C25, C33, C56	0	1	1	1	1	0.08	0.78
9	C1, C17, C42, C46	1	0	0	1	1	0.04	0.76
10	C18, C29, C30, C43, C79	1	0	1	0	1	0.06	0.75
11	C5, C37, C62, C64, C77, C78, C81, C86	1	0	1	1	1	0.09	0.78
12	C13, C23, C27, C54, C59, C69, C74, C80, C83, C84, C89	1	1	0	0	1	0.12	0.76
13	C2, C15, C19, C50, C68, C71, C73, C76, C88	1	1	0	1	1	0.10	0.79
14	C3, C4, C9, C16, C20, C22, C31, C32, C35, C45, C57, C65, C67, C85, C87	1	1	1	0	1	0.17	0.78

	Case ID	Schema share effectiveness	Agent network	Problem solving effectiveness	System signals	Outcome	Coverage	BN Prediction of Success
15	C7, C8, C24, C26, C38, C40, C41, C51, C70, C75	1	1	1	1	1	0.11	0.81

Table 9.2 summarises the increasing probability of success against an increasing number of core CQIN constructs being True.

Table 9.2: Success Range Based on Number of Core Constructs = True

Number of core constructs = True	Probability of success
0	0.17
1	0.59 - 0.66
2	0.71 - 0.76
3	0.78 - 0.79
4	0.81

The results of the CQIN BN are consistent with both the ABM results and the secondary data analysis. Substantial additional insight is provided by the BN representation of CQIN. The problem of low resolution and sparse data (missing observations) is mitigated. Alternative hypotheses for variable interaction (using OR or AND logical functions) can be applied, and the notion of probability distributions as applied to QI is illustrated. What does it mean for a QI initiative to have a 71% probability of success? Intuitively this seems very different to an 11% or 95% probability, but how should QI practitioners respond to this information?

9.8. Critical Assessment of the BN Analysis

A structured development process was followed consistent with the recommendations of Kyrimi et al., (2021): (i) the purpose of the BN is provided – to model the probability of success for a QI within CAS initiative; (ii) all the variables and the network structure have

been described; (iii) the node probability tables are described, including (iv) the role of empirical data in forming the expression for the node probability tables, and (v) initial internal validation was conducted. Wider external validation and using the model in practice are the subject of ongoing research.

The limitations of the data used to construct the CQIN BN were recognised and mitigated. The result is a transparent model that can assess the probability of success for QI within CAS. Having reached this stage of development, ongoing refinements can be made as more empirical data is available. The relative simplicity of the variables and structure make collection of empirical data a realistic task, whether this is via further secondary data review or in-depth observed case studies.

Chapter 10 Insights from the Applied Case Study

10.1. Introduction

This chapter describes the applied case study undertaken within the research. The participating organisation is introduced in section 10.2 and details of the observed QI project are set out in section 10.3. The engagement process and activities are covered in section 10.4 and section 10.5 presents the quantitative results. This includes the process of translating the case study feedback into a face validity assessment for the agent-based model. Section 10.6 presents qualitative insights gleaned from the case study, specifically on agent networks and problem solving. The chapter concludes with a summary of the case study in the context of this research (section 10.7). Section 10.8 is a brief coda to report the progress and the current status of the observed QI project.

10.2. Organisational Background

The case study organisation is a large private provider of aged residential care (ARC) services in New Zealand and Australia. This organisation is listed on the New Zealand stock exchange and operates multiple retirement villages in New Zealand and Australia. In the New Zealand ARC sector, large commercial providers operate a mixed model of retirement living and healthcare provision. As a major provider of ARC in New Zealand, the case study organisation invests in a data-driven approach to care and actively supports aged care research.

Within its clinical operational activity, the organisation maintains an interdisciplinary medications advisory committee (MAC) that oversees performance and improvement of medication-related care. The committee covers clinical and operational functions and includes general practice representatives, a pharmacist, geriatrician, specialist researcher, nurses, and improvement support staff. A priority project for the MAC is the reduction of prescribing antipsychotic medication when not clinically indicated, a major focus for care of residents living with dementia.

10.3. The Improvement Project

A significant proportion of people living in residential aged care are exposed to antipsychotics, with approximately one-third of residents receiving a medicine in this class at some point during their admission. Antipsychotic use is associated with significant adverse effects and an increased risk of premature death, especially for people

with dementia. Studies of antipsychotic deprescribing in people living in residential care have produced mixed results (Van Leeuwen et al., 2018). Some trials have successfully withdrawn antipsychotics in participants without apparent adverse effect, but other trials have either failed to reduce antipsychotic use successfully, failed to show any benefit associated with cessation, or have reported an increase in adverse outcomes associated with deprescribing (Birkenhäger-Gillesse et al., 2018).

Non-pharmacological management strategies are the preferred alternative to antipsychotics for managing distressed behaviour in people with dementia, but while these strategies reduce the need for pharmacological and physical restraints, they are not a panacea (Birkenhäger-Gillesse et al., 2018). Inadequately controlled aggressive or agitated behaviour creates physical risk for care staff, other residents, and the affected individual. Other behavioural symptoms of dementia can cause stress and burn-out in carers, reducing their capacity to provide effective person-centred and compassionate care.

When initiating an antipsychotic for people with behavioural and psychological symptoms of dementia, prescribers must consider the potential benefits and harms of *using* antipsychotics along with the potential benefits and harms of *not using* them. When evaluating deprescribing of antipsychotics, similar consideration is given to both arms (*use* versus *not use*) of this harms and benefits dichotomy. There is no simple guideline or one-size-fits-all solution to this complex issue. Antipsychotic deprescribing is an ideal initial verification case for CQIN because it is so complex. Minimising potentially clinically unnecessary prescribing involves a large heterogenous network of agents across multiple organisations and professional disciplines.

The interventional approach for this QI project is premised on providing timely, targeted prescribing data to prescribers and encouraging discussion and consideration of alternatives. The underlying hypothesis is that ensuring the appropriate information is readily available to those making patient-level decisions, combined with raised awareness of the alternatives, will allow front-line prescribers to identify early where any reduction in antipsychotics may be appropriate.

10.4. Engagement

Engagement began by introducing MAC members to the CQIN conceptual model and discussing the antipsychotic activity in terms of CQIN. The author had the role of invited

observer of the deprescribing project, able to participate in regular reviews and ask questions of individual MAC members. The MAC team understood they were assisting with a research study and viewed this as an opportunity to look at their QI practices from a new perspective. After establishing the context and describing the roles and interactions involved in the project, the team considered plausible starting values for the simulation model variables such as *system signals* and *problem solving effectiveness*. This was not a trivial task given the novelty and abstract nature of CQIN. Clinicians and improvement facilitators were thoroughly versed in the clinical and change management details, but CQIN focuses on shared understanding, network relationships, and problem-solving activity. The author worked with small subsets of the medications advisory committee to first introduce and illustrate the concepts and then consider potential values for the CQIN variables in the context of the deprescribing project. Obtaining valid expert assessments of probability is recognised as difficult and error-prone (Fenton & Neil 2019; Morgan 2014). An elicitation protocol was developed and the information gathering process was highly iterative, updating as new information or understanding became available.

After multiple rounds of elicitation activity, including an anonymous survey, a credible set of starting values for simulation was established. These values were used to form the basis of the starting conditions for the ABM simulation experiments already introduced (Table 7.1, Chapter 7). Appendix 6 contains the survey questions used to collect each participant's interpretation of their project in terms of the CQIN model. As engagement with the MAC progressed, the focus narrowed to: (i) examining the role of system signals; (ii) interpreting the agent network; (iii) the level of shared understanding, and (iv) the problem-solving options. This focus emerged naturally from conversations with the participants, however it is noteworthy that these are same four positive core CQIN constructs as identified within the secondary data thematic analysis.

10.5. Results

Half of the 12 MAC members completed the anonymous survey. The objectives for the survey were to establish realistic representation of the deprescribing project for simulation and to better understand how improvement practitioners interpreted the CQIN constructs. Table 10.1 sets out how the combination of ranked scores and qualitative comments were translated to starting conditions for the simulation experiments. For the ranked scores, the minimum and maximum scores are presented to identify the range of opinions held by the participants.

Table 10.1: Translation of Survey Results into Model Parameters

Survey Questions	**Min Score**	**Max Score**	**Is an indicator for (in plain English)**	**Simulation Model Parameter**
[Core & Related Network questions]	na	na	Ease of sharing information and co-ordinating actions across multiple participating groups and individuals	Minimum node degree of agent network
Improvement outcome feedback strength	2	5	The availability and quality of feedback on your process	Signal detection level
Other system feedback strength	2	5		
Team alignment of project goals	3	5	Shared purpose; agreement as to the 'what, why and how'	Schema share effectiveness
Team alignment of project interventions	3	5		
Problem solving effectiveness	3	4	Selection of the right interventions and strategies for this problem, Building knowledge from each intervention/trial	Problem solving effectiveness Learning gain
[Free-text descriptions of Problem-solving approach]	na	na		
Success prediction	3	5	Probability of the project achieving its goals	Target outcome

Numerical Scale Scoring:
1 = Very Weak; 2 = Weak; 3 = Moderate; 4 = Strong; 5 = Very Strong

The results in Table 19 show that this group of participants held a range of subjective assessments as to the strength of key CQIN indicators. Although expected and perhaps obvious in hindsight, it was a valuable illustration that alignment of interpretations cannot

be assumed. Additionally, having a range of estimates requires steps to ensure reliability of measurement, to minimise interpretation errors and preserve genuine variation. For this survey, participants had two previous sessions to become familiar with CQIN concepts and terminology, and each survey question was primed with a definition as in the example below for distinguishing between core and related agent networks:

We describe related networks as being needed to support or contribute to the changes you want to make, but they are not part of your regular information sharing and decision making. If an agent group is part of your regular information sharing and decision making, they are part of your core improvement network.

The survey responses revealed a range from 60% to 100% for team alignment of project goals, project interventions, problem solving effectiveness and probability of overall success. For improvement outcomes and other system feedback, the minimum assessment dropped to 20%. The range in responses was presented back to the MAC team together with a simulation example. As detailed in Chapter 7, the baseline simulation experiment returned a probability of success for the project within 65% and 69%, depending on the number of replications. The simulation result was thus at the lower end of the respondents' success predictions (from 60% to 100%).

10.6. Qualitative Interpretation

The case study provided feedback on the interpretation of CQIN constructs and the potential utility of CQIN as a diagnostic tool for complex quality improvement initiatives. Free text responses within the survey and follow up interviews added details, particularly to the understanding of agent networks and problem solving.

10.6.1. Agent Networks

The variation in the interpretation of agent networks is shown in Table 10.2. It is emphasised that these respondents had been familiarised with the concept of an agent network and were now trying to relate this concept to their practical experience. For the antipsychotic deprescribing project, professional roles and facilities seemed to be interpreted as boundaries, with a 'network of networks' emerging across general practitioners, nurse managers, nurses, carers, facility managers, pharmacists, patients, and family members. These professional boundaries were repeated in a modular fashion across facilities. Additionally, the level of engagement within a role or facility also influenced the interpretation.

Table 10.2: Interpretations of Core and Related Agent Networks

Agent involvement	**Assignment as Core vs Related**	**Variations in interpretation**
1) Active contributors to the improvement objectives, decisions, and clinical processes	Core network	Consistently interpreted
2) Nominally in core agent network but variation in levels of individual participation	Initially assumed to be core; some interpretation as related	Interpretation being made at individual rather than group level
3) Active contributors to the improvement process and/or clinical care delivery but viewed as outside the *current decision making and information sharing*	Initially assumed to be core; some interpretation as related	Not in decision - making network
4) Contributors to improvement process but viewed as outside the *clinical care decisions and delivery*	Initially assumed to be core; some interpretation as related	'Core' was interpreted as the people who are actually involved in the delivery of care
5) Aware of and support the activity but not actively involved	Related network	Consistently interpreted

As labels, the distinctions between core/related network are not critical. However, the varied interpretation of multiple 'teams within teams' is an important finding. Increased involvement in information sharing and decision making is theorised to be essential to purposeful change within healthcare (Damschroder et al., 2009). In practice this means paying attention at an individual level (change champions acting as highly connected and influential network nodes) and at the team level (ensuring multiple connection paths into teams and shared understanding within key subgroups). 'Teams of Teams' are vulnerable to network disruption if critical nodes (key individuals) control information and then drop out of the wider network (Barabási, 2016). This effect was

observed to a small extent within the project, where some MAC members perceived their ability to contribute and the level of information sharing as less than ideal. In some cases, barriers between facilities and professional disciplines were identified. As a contribution to the present research, the case study interpretation of agent networks reinforces that the definition of an agent network is fundamentally subjective and will inevitably be interpreted differently by different observers. This finding is consistent with the literature on agent networks (Nair & Reed-Tschocas, 2019). It is also consistent with the secondary data analysis, where strengthening and growing network connections was identified as an essential strategy, without a fixed interpretation of network boundaries and no discernible focus on formally defining them.

10.6.2. Problem Solving

Two survey questions deliberately avoided any CQIN-related or complex system jargon. These questions returned a wide range of responses, and by implication, variation in the respondent's priorities for the project.

> *What do you think is the most important improvement action to take at this stage of your project?*

Paraphrased answers:

- Find out more precisely what works and why
- Strengthen the feedback loops
- Design a scaled-up intervention (2 responses)
- Reinforce the "Why" of the project
- Take a closer look at clinical distinctions between medications

> *What do you consider to be the major constraint or obstacle to the project success?*

Paraphrased answers:

- Lack of available time to focus on this initiative (2 responses)
- External pressures e.g. COVID, wider health system demands
- Lack of engagement at village (facility) staff level (2 responses)
- Hard to maintain momentum
- Geographical spread; difficult to maintain personal touch

These two questions explored the shared understanding of problem-solving strategies amongst the MAC members. The answers above indicate the depth of problem-solving knowledge present within the group, and strategies for future action. To an extent these responses surprised the group, who considered themselves well-aligned on what to do next. Specifically, clarity of the change hypothesis and how to progress beyond a pilot phase were questioned. The answers provided above suggested multiple potential hypotheses within the group as to what to do next and why. The range in responses illustrated the value of pragmatically examining the schema share effectiveness and problem-solving effectiveness within an agent network.

10.7. Summary of the Applied Case Study

The applied case study made a small but essential contribution to this research. It was an opportunity to test the theoretical concepts of CQIN with a real-world complex quality improvement project. Perhaps unusually for a case study, the research objective was not to generate quantitative data or observational insights for theory-building. Instead, the case study functioned as a proof of concept and valuable experience for the author in explaining abstract constructs such as *schema share effectiveness* or *system signals.* Important nuance was added to the interpretation of agent networks and problem-solving approaches. The project also demonstrated the difficulty in developing a practical framework, with the need to carefully navigate team dynamics, interpretive biases and establish the reliability of aggregated views. The end results were able to be used for the simulation experiments and were consistent with the secondary data analysis of other QI projects. Positive feedback was received from MAC members following the presentation of a summary report, particularly with the emphasis CQIN placed on information sharing across a large and heterogenous network.

10.8. Real-World Progress of the Antipsychotic Deprescribing Project

The deprescribing initiative continues as a long-term project for the case study organisation. Within the pilot facility, the proportion of residents with antipsychotic medication administered reduced from a peak of 77.5% in June 2021 to 45.2% in May 2023. For the same time period, the baseline level of prescribing also reduced, but not as dramatically – from 87.5% to 64.3%. Therefore, there was a reduction of antipsychotic prescribing of 32.3% against a baseline reduction in prescribing of 23.2%.

In the framing of CQIN, the problem-solving strategy has been to provide relevant, timely feedback signals to prescribers. The expectation is that this encourages different clinical behaviour, whilst still leaving clinicians autonomy to manage inevitable patient variation and context within each prescribing instance. Fundamental to the strategy is the availability on non-pharmacological care models and here the organisation has, in CQIN terms, a major coevolution constraint. Non-pharmacological strategies to support residents with dementia are labour intensive and take time to establish. At the time of writing, the resources are not yet in place to support rapid replication of the pilot facility success.

Part 3: Discussion and Conclusion

Chapter 11 Discussion

11.1. Introduction

This chapter looks at the results presented in the preceding chapters and considers their collective impact. This includes the overall picture formed and the contribution of each research component. A critical review of the results is presented first (section 11.2), followed in section 11.3 by a summary of six key findings that inform the subsequent discussion. An evaluation of the research hypotheses given these results is provided in section 11.4, leading naturally on to an assessment of the research question in section 11.5. Section 11.6 contains a detailed consideration of the CAS explanation of QI and the foundations for this within an evolutionary process. Implications for theory and practical applications for improvement practice follow as sections 11.7 and 11.8 respectively. Section 11.9 restates the key findings, implications and potential opportunities.

11.2. Critical Review of Results

A conceptual model for QI within a socio-technical CAS was developed from relevant literature. This model was then operationalised as an agent-based simulation model, with insights from an applied case study refining the constructs and serving as a face validity example. Secondary case data from published QI projects was examined, indicating empirical support for the model and providing additional details. Finally, each of these research strands was brought together into a Bayesian network model representing the structural relationships and interactions of QI activity interpreted as a CAS. Considered as a whole, the research threads build a detailed and consistent picture. Triangulation of multiple data sources was achieved via the QI literature, applied case study, CAS theory and simulation experiment results.

Each component of the research had constraints. Efforts were made to identify these issues before commencing and to put mitigations in place. The nature of complex systems introduced specific challenges. A valid criticism of the simulation experiments is that they were non-ergodic, in the sense that they only explored in detail a small number of potential system configurations from within the set of all the possible configurations. In response, it is noted that the sensitivity analysis revealed a narrow operating range for the model, with combinations outside this range offering little new information beyond a guaranteed failure/guaranteed succeed outcome. The exact range of operation is a direct reflection of the modelling parameters, and it is not claimed that the modelled range

exactly reflects real-world settings. No attempt to smooth or weight individual variables was made in this iteration of simulation experiments. Nevertheless, the ability to recreate a bounded operational range was a useful experimental finding. The concept of a valid operating range for system variable combinations is consistent with CAS principles (Ashby, 1956; Maturana & Varela, 1987; Mitchell, 2009).

Ideally, more opportunities to work alongside and directly observe QI teams in an applied manner would have been included in this study. In-depth qualitative observation was unrealistic for two main reasons: (i) the effort required to sufficiently introduce and operationalise abstract complexity concepts, and (ii) external pressures on the target workforce groups. The use of secondary data allowed for an empirical assessment of the conceptual model but introduced new limitations including publication bias and incomplete reporting of relevant factors. Potentially the most significant challenge was the lack of consistent operational definitions across the cases. Consistent with the reviews of Thompson et al. (2016) and Turner and Baker (2019), the usage of complexity terms was fragmented and only rarely integrated into a theory or model.

An objective measure of 'successful' QI activity was not possible. Self-reported success was accepted at face value and recognised as a moment-in-time outcome, with no assertions of long-term sustainability. However it is noted that there was a small number of cases (7) that did report ongoing sustainability, describing actions over longer timeframes (Brys et al., 2019; Carter et al., 2017; Clancy, 2009; Czulada, 2016; Dandoy et al. 2015; Ngo et al., 2022; Poksińska, & Swartling, 2018).

The obvious issue for the QCA was the very unbalanced data set. Recognising this limitation, the necessity of any given construct or set of constructs is not claimed. To assess the 'if NOT Y, then NOT X' criterion of a necessary condition, more cases are required where the reported outcome was unsuccessful, and the construct of interest was not present (Ragin, 1999). The 'if X, then Y' criterion for a sufficient condition can be more confidently stated, accepting the self-reported definitions of success. The QCA identified multiple sets of sufficient conditions consistently reported across 90 QI cases. Treating the reported success factors as sufficient sets (i.e. *this* combination of factors was successful in *this* instance) overcomes one of the principal difficulties with complex system research. Each case remains holistic with a unique context, but common patterns can be identified across the many different contexts.

11.3. Six Important Findings

As a pre-requisite to assessing the research hypotheses and the theoretical implications of the research, six important findings of the research are presented first. These findings are a mixture of hypothesised system behaviour, unexpected insights and useful methodological strategies. They form the building blocks for a CAS explanation of QI.

11.3.1. Non-linear Effects

The results of the simulation experiments displayed the general CAS principles of uncertainty, non-linearity, dependence on starting values, and inter-dependence with other system variables. Experiment 2 demonstrated a clear threshold "tipping point" for the schema share effectiveness variable, and Experiment 3 revealed further non-linear sensitivity of critical parameters such as the minimum cooperation threshold and agent node degree. The non-linearity within the QI activity arises directly from the three 'stop' parameters within the model – either the improvement target is achieved, the minimum cooperation level is not maintained, or the planned number of iterations is completed (i.e. a predetermined investment of effort is reached).

Variations in the agent network structure, together with the uncertainty of success associated with any agent interaction, create conditions where cascading failure is possible (Mitchell, 2009). A large, heterogenous agent network is made up of multiple communities and clusters with different levels of connection (Barabási, 2016). In these conditions, individual agents with high node degree who stop cooperating have a large impact and can trigger a cascade of network losses. A genuinely complex problem means that shared problem interpretation and physical cooperation are difficult to achieve, and this results in a higher probability of resistance. These network losses lead to the phase transitions observed in the simulations. The transition between states is a form of self-organised criticality, manifesting as a cascading failure triggered from the expected micro-level system activity (Bak & Chen, 1991). For the deprescribing project simulations, a minimum level of cooperation was set as 50% of the network. Accepting that the location of the resulting transitions is not empirically precise, it is still taken as axiomatic that *any* agent network has a minimum critical mass for cooperation to be sustained. System state transitions have significant implications for QI practice. Working with a CAS explanation of change, it becomes essential to identify the minimum values of the critical CAS variables and also where the improvement network sits in relation to the system transition points.

11.3.2. Minimum Viability Levels of System Constructs

Conceiving system constructs as having thresholds or bounded ranges of effectiveness creates a new emphasis for empirical measurement of construct values. Rather than asking *what is the level of team problem solving* using a validated scale for team problem solving, the critical question to answer is *has our problem solving achieved a threshold to be effective in this specific context?* Conversely, teams need to know if a given construct has fallen below the effectiveness threshold required for the given context. QI practitioners can legitimately target at-risk or below threshold variables. Similarly, QI-informed practitioners would recognise that if important constructs are below their minimum threshold for effectiveness, focusing attention on other variables will be of limited value.

11.3.3. Learning Gain

Learning is conventionally described as an outcome of applied QI activity (Lemieux-Charles et al., 2002). To respond to system signals, an agent network must rely on the levels of knowledge already within the network, or else generate new knowledge. New knowledge can be learned empirically, using problem-solving techniques such as trial and error, means-end analysis, or hypothesis testing (Anderson, 2005). Knowledge can also be imported by extending the network and bringing in more agents with different knowledge.

An agent network can therefore maintain a range of desired outcomes using existing knowledge and creating new knowledge when required. However, as complexity levels increase, it is likely that the level of new knowledge required is also high, and the knowledge content required may also change rapidly. The learning process must be able to generate many new knowledge elements. These elements can then be combined to form building blocks that lead to increasingly complex solutions not previously conceived or achievable within the learning process (Dennett, 2017). Therefore, learning in complex systems will be highly iterative, with previously acquired knowledge re-invested as an *input* into the next learning cycle. Without this ongoing positive learning gain an agent network will only be able to respond to dynamic system changes using a finite and static set of options. QI within CAS requires the ability to learn quickly and grow knowledge to successfully adapt.

11.3.4. Sets of Sufficient Conditions

A methodological challenge in studying complex systems is to retain sufficient context to see the 'system' in its entirety, yet still have a granular focus to examine the detailed interactions between variables. With high numbers of variables and potential interactions, conceptually representing these structures is difficult. Causal loop diagrams are one way of achieving this, but this approach relies on some understanding of structural connections that is not realistic to reproduce using the limited contextual information supplied in the QI case study literature.

Identifying detailed causal pathways is difficult when the structure is hard to comprehend. The configurational approach to causality of QCA allowed associations to be established across multiple cases with very different subject matter and wider system contexts. The differing combinations were minimised to logically distinct combinations of variable states — 'solution sets' as the unit of analysis, not an isolated variable. The solution sets revealed common patterns and recurring emphasis, which can be thought of as *strategies* given a system context. 90 disparate QI case studies were able to be mapped to 15 solution sets, and further minimised to 3 solution strategies.

11.3.5. Directions of Influence Within a Conditional Probability Network

The solution sets were not 'structure-free' correlations. Causal assumptions are embedded in scientific reasoning and representation in some form of path diagram makes these assumptions explicit and testable (Pearl & Mackenzie, 2018). Directional network paths were identified from fundamental reasoning idioms, such as cause-consequence, where B follows A, but A does not follow B. Using definitional reasoning, variables were described at a high level to facilitate analysis, with the ability to deconstruct into more granular constituent parts as needed.

Empirical evaluation, via the secondary data results, supported the hypothesised variables and their proposed interaction. It was necessary to include one further layer of granularity into the original CQIN model, but the core and contributory levels proposed matched the data and no obviously missing variables were identified. Proposed connections were generally supported, but two direct paths from the *system signals* and *agent network communication* variables to the outcome variable were suggested by the data. The subsequent Bayesian network representation was also consistent with the hypothesised causality assumptions. Minor connections are almost certainly missing from

the model for example, *problem space complexity* could be reasonably proposed to influence *schema share effectiveness*. In the Bayesian network context this relationship is still present as both nodes are parent nodes with a shared child node and are therefore not fully independent. Representation of QI as a probability network offers an alternative modelling approach when exhaustive empirical observation and measurement is not feasible (Fenton & Neil, 2019; Pearl & Mackenzie, 2018).

11.3.6. Mediation and Moderation Effects

The Bayesian network representation of QI facilitates the analysis of how each variable contributes to the outcome. Immediately apparent when looking at the CQIN BN model is the marginal impact of isolated interventions on single variables. Serial connection paths identify mediation effects, for example, improving *interpersonal factors* influences the outcome via the increase in *schema share effectiveness.* Importantly, as the length of serial paths increases, the ability of any one variable to contribute on its own diminishes. For example *autonomy* is identified as important to QI activity, but improving *autonomy* in isolation only has a limited impact on *problem solving effectiveness*. Other parent nodes also moderate *problem solving effectiveness*, and then the impact of *problem solving effectiveness* on the outcome is itself moderated by the other parent nodes. The originally hypothesised moderation impact of *coevolution constraints* and *problem space complexity* is accounted for, with these variables also being direct parents of the outcome variable. Positive effects from any other parent nodes must overcome any inhibiting effects from these moderators.

This structure of core and contributory levels with parallel parent-child paths results in a highly generalisable analysis tool. The impacts of specific QI interventions can be explored prior to undertaking potentially disruptive change and the genuinely influential leverage points within a given QI context can be identified.

Having highlighted the important findings, section 11.4 now returns to the research hypotheses, and evaluates the answers this study has provided.

11.4. Research Hypotheses

11.4.1. Hypotheses 1 and 2

Hypotheses 1-2 served as preliminary assumptions for the research question.

H1. Change within CAS will be consistent with CAS principles.

H2. Conditions for successful change within CAS can be identified.

As detailed in Chapters 1 and 2, a substantial body of literature makes the claim that complex human socio-technical structures do function as CAS. Accepting that stance as the established position within the discourse, H1 naturally follows. If socio-technical CAS are a legitimate subgroup of general CAS, then CAS principles should apply. The limited resistance to the notion of healthcare systems as CAS is focused on H1, suggesting human agency and goal awareness contradict self-organisation and the emergence of system-level outcomes from simple, predetermined rules (Paley & Eva, 2011). Evidence to support H1 address this research gap and has wider implications for improvement practitioners.

H1 was supported by this study. The results show that a purely CAS explanation of change within complex socio-technical contexts is feasible. The *schema share effectiveness* construct represents how the goals, assumptions, and problem-solving options form a foundation of shared problem representation amongst agents (Anderson, 2005; Cronin & Weingart, 2007, 2019; Holland, 2012; Mohammed et al., 2010). In the simulation experiments, the process of responding to a signal for change, acquiring sufficient knowledge and then cooperating to achieve change was achieved exclusively through the five information-sharing and cooperation rules. No other mechanism for aligning agent schema existed within the simulation. *Problem solving* provides the cognitive process equivalent to the biological processes taking place in natural CAS. Importantly, the share/cooperate rules on their own are not sufficient to ensure success. *New* knowledge must be accessed or generated, selected, and applied in real-world tests of fitness. At the agent level, adaptation is the updating of the internal schema. If system conditions allow, system-level adaptation takes place when the level of aligned schema across the agent network is sufficient to enact change and successfully reach a goal state (Carmichael & Hadžikadić, 2019). An individual's micro-level agency ultimately has consequences that effect the system, and these effects manifest at the macro (system) level (Carmichael & Hadžikadić. 2019; Wilson & Kirman 2016).

Strong support for H1 was also evident in the secondary data analysis, via the consistency and coverage of the CQIN constructs across the cases and the absence of themes in conflict with H2. The role of leadership could potentially be viewed as one example contradicting CAS principles, and the implications of this are examined in sections 11.5 and 11.6.

H2 posits that conditions for successful change within CAS can be identified. This hypothesis was supported across all the research strands. Both the agent-based simulation and the secondary data analysis identified influential variables. Different combinations of variables across different contexts form sets of sufficient conditions for successful change. Section 11.5 presents the detailed results for each of the original CQIN variables and the additional findings.

11.4.2. Hypotheses 3 and 4

As noted in Chapter 1, H1 and H2 are broad statements, containing multiple CAS principles which might be tested in different ways. H3 considers the role of the agent internal schema in generating emergent, system-level QI phenomena. Closing representation gaps amongst agents was proposed as the spreading mechanism across the problem-solving network (Cronin & Weingart, 2007). H4 bounds any alignment of agent schema to a limited set of agent behavioural rules, a core principle of CAS.

H3. Within socio-technical CAS, system-level change in the form of goal-directed improvement is achieved through successfully coordinated updates and alignment of agent schema.

H4 Updating and aligning agent schema is achieved within the constraints of a limited set of agent behavioural rules.

The agent-based simulation demonstrated that H3 is plausible, with nominal QI targets able to be achieved by accumulating new knowledge via repeated learning cycles updating the agent internal schema. The number of updates required, and the amount of alignment (cooperation) required within the network are treated as context-specific. In other words, the number of learning cycles and the level of cooperation required will be determined by the improvement problem being addressed.

The agent-based simulation modelling also demonstrated that a small, finite set of behavioural rules can lead to system-level change (H4). Key to this understanding in a

socio-technical CAS is the rules being expressed at an abstract level. *Cooperate* or *don't cooperate* are simple rules, but these will manifest with great variety in real-world improvement activity. These rules describe the *impact* or *outcome* of the granular agent-level interaction, not the specific actions being taking place within clinical and supporting cooperation. Under conditions of complexity, a desired QI outcome is unlikely to be observed as a direct, linear outcome of specific interventions. CQIN proposes that sufficient cooperation can lead to the desired QI outcomes. In CAS language, QI emerges from the lower-level intervention activity. Are the conditions (simple CAS rules) for emergent phenomena present? The simulation modelling and secondary data analysis both supported H4. Attention to a conceptual level that assesses the conditions for emergence therefore seems essential to a CAS interpretation of QI.

11.5. Research Question

What system factors influence the success or failure of healthcare QI activity within complex adaptive systems?

The results for the 6 core CQIN constructs are now considered in the context of current understanding and previous research. The definition for each construct is restated first in Table 11.1, along with any development of this construct that took place across the research phases (reading Table 11.1 from left to right). The letters in brackets refer to the original construct labels as introduced in the conceptual model. Changes over time reflect refinements in labelling and further definition of sub-elements when required by each research phase. For example, the parameters required to run the ABM, and themes within the secondary data analysis consistently described at a different level of granularity than the original definition.

Table 11.1: Evolution of the CQIN Constructs

Conceptual Model		Agent-Based Model		Thematic Analysis and Bayesian Network	
CQIN Construct	**Definition**	**CQIN Construct**	**Configurable Parameters in ABM**	**CQIN Construct (Core)**	**Parent (Contributory) Nodes**
Agent network (A)	The agent network is made up of all the participants and affected parties of the QI activity. For healthcare QI activity, the agent network encompasses clinicians, patients, QI practitioners, management and supporting functions, within the interpreted environmental boundaries.	Agent network (A)	Number of networked agents (nodes)	Agent network communication paths (A)	
			Network type (preferential attachment or random)		
			Minimum node degree		
			Network centrality measures		
			Proportion of network cooperating		

Conceptual Model		Agent-Based Model		Thematic Analysis and Bayesian Network	
			Network rejection scope (range at risk)		
			Network rejection probability		
			Minimum cooperation threshold		
Operational processes (B)	Operational processes and activity as relevant to the improvement context, for example – processes and applicable clinical protocols in a given healthcare context.	Operational processes and Outcomes (B), (C)	Outcome measure at start	Target improvement level achieved (B, C)	
Operational outcomes (C)	Defined, measurable process and system outcomes.		Targeted outcome measure		
			Elapsed time		

Conceptual Model		Agent-Based Model		Thematic Analysis and Bayesian Network	
System signals (D), (E)	Feedback on the health outcomes (i.e. changes in specified outcome measures) as well as wider system feedback (such as changes in the environment, resources, process timeliness or costs).	System signals (D), (E)	The range of initial signal detection amongst agents	System signals (D, E)	
Shared schema (F)	The level of coherently aligned goals, assumptions, priorities, and improvement options across the agent network.	Shared schema (F)	Schema share effectiveness	Schema share effectiveness (F)	Interpersonal Factors
					Learning gain
Problem solving (G)	The domain knowledge, problem-solving strategies, and resources (time, funding, cognitive bandwidth) available to the network.	Problem solving (G)	Problem-solving effectiveness	Problem solving effectiveness (G)	Autonomy
					Option Generation
					Focus / Time
					Learning gain
					Budget

Conceptual Model		Agent-Based Model		Thematic Analysis and Bayesian Network	
System complexity factors (H)	The problem space is measured by the system complexity factors present. Using Wiesner & Ladyman's (2019) conditions for complexity, these are: (i) numerosity of elements; (ii) numerosity of interactions; (iii) diversity of elements; (iv) non-equilibrium; and (v) feedback. Products (outcomes) of complexity are: (i) non-linearity; (ii) self-organisation; (iii) adaptive behaviour, and (iv) memory.	System complexity factors (H)	Problem space complexity	Problem space complexity (H)	

Conceptual Model		Agent-Based Model		Thematic Analysis and Bayesian Network	
External environment (I)	The external environment is that part of the contextual setting that falls outside the interpreted and enacted boundary (Nair & Reed-Tsochas, 2019). As this boundary is dynamic and may vary between agents, it acts as a moderating factor, in the form of coevolution constraints (the impact from shared and adjacent networks).	External environment (I)	External (coevolution) constraints	Coevolution constraints (I)	
Learning gain	The iterative nature of CQIN is difficult to convey. CQIN functions as a series of learning cycles. A learning gain from each iteration is posited as critical to acquiring sufficient knowledge within the agent network (Deming 1986; Doolittle 2014; Reed et al., 2018).	Learning	QI phase (e.g. plan or do) Learning gain per cycle Improvement increment count Number of planned iterations	(Mapped as a parent node of both schema share effectiveness and problem solving effectiveness)	

11.5.1. Schema share effectiveness

The schema share effectiveness construct measures the ability to spread or replicate a mental model of the problem and desired actions. It is emphasised again that this is not meant to imply an exact copy, where all agents now have the same mental model. A shared problem representation means that within two unique agent worldviews, a given signal is noticed, and appropriate next steps are agreed. Schema share effectiveness thus measures cooperation and the level of spread of an idea. It is also assumed that agreement is unlikely to occur in one simple transaction. Multiple opportunities to align schemas are required, as are enabling factors such as common experience and trust. Within the secondary data analysis, 69% of cases identified actions to increase schema share effectiveness directly or indirectly. This result supports the originally hypothesised importance of this construct based on the literature on mental models (Cronin & Weingart, 2007, 2019; Filho & Tenenbaum, 2020; Wang et al., 2016). It is consistent with the 'disciplines' of shared vision and team learning within organisational learning (Senge, 2006). The requirement to effectively align and share agent schema, as a necessary component of change, is further supported by empirical examples of the critical role of conversation and communication in healthcare improvement (Blackman et al., 2022; Cribb et al., 2022).

11.5.2. Agent network

The agent network provides the substrate for interactions and communication amongst agents. This construct should perhaps be broken down further in subsequent research. For this study, 'agent network' and 'agent network communication paths' are used interchangeably. In other words, it is the connections between agents being represented, not the individual agent characteristics. However, personal knowledge and skills have a role within the schema share effectiveness and problem solving effectiveness constructs. 66% of the secondary data cases referred to increasing communication and opportunities for communication, across different groups of agents.

The applied case study revealed the highly subjective nature of determining agent boundaries. Each agent will interpret their boundaries differently and the structure is not fixed in time. In healthcare QI, the role of patients as agents and the concept of 'cocreating healthcare' was prominent in the secondary data (Bishop & Waring, 2019; Hignett et al., 2015; Lachman et al., 2015; Myron et al., 2017; Yardley et al., 2022).

Studying the agent networks in greater detail is a rich vein for further QI as CAS research. How is self-organisation reflected across small teams? What are effective strategies for improving the 'opportunities for communication'? How do the network structures align with organisational boundaries and hierarchies? Finally, it is noted that the hypothesised function of agent network 'viability' was not observed directly in the applied case nor the secondary data. Nevertheless, some minimum level of agent cooperation is assumed, below which effective change will not occur. Given the strong influence of schema share effectiveness and the agent network communication paths, researching ways to measure them accurately will be invaluable.

11.5.3. Problem solving effectiveness

Problem solving effectiveness was observed as a success factor in 66% of the secondary data cases. This construct covers the strategies, methods and tools used to resolve problems, so prominence as a success factor is unsurprising. More informative, as the QCA revealed, problem solving effectiveness is only one of four potential solution focus areas. It is also more likely to be effective *in combination* with the others. As with shared understanding, the construct measures the impact of the problem-solving activity, the 'effectiveness' will depend on choice of methods, skill in application, and the difficulty of the problem. Within the secondary data, there was little reporting of how problem-solving strategies were considered and selected. Usually, the chosen approach such as Lean, PDSA, or healthcare collaboratives was presented and justified. This pattern no doubt reflects the nature of reporting empirical case studies. However, the applied case study revealed that in practice, improvement teams may not be fully aligned as to the preferred problem-solving strategy. Ongoing confirmation that a given approach is working, and consideration of alternatives, seems important.

Deconstructed one level, problem solving was separated into capability, comprised of skills and resources (capacity). Skills included option generation and the ability to learn. Resources included autonomy to make decisions, time and focus to consider a problem, and budget. This level of definition was adequate to analyse the secondary case data, but one additionally important problem-solving factor was also identified – professional capability. Unsurprisingly, experienced doctors, nurses and allied health professionals are more likely to be able to resolve issues. Therefore, professional development is itself an effective QI strategy, provided these highly skilled agents have

the time and autonomy to attend to issues and train others. Specific QI training will augment, but not replace, professional expertise.

11.5.4. System signals

The last of the four core constructs theorised as positively influencing improvement outcomes, system signals were identified as a success factor in 51% of the secondary data cases. Again, this is an abstract construct that can be broken down to include such items as signal detection, signal timing, signal strength, and signal validity. System signals represent the information circulating amongst the agents, via the network communication paths. An unanticipated finding from the secondary data was that attending to system signals could be an effective change intervention in its own right (Alhaider et al., 2020; Elertson, 2017; Murray et al., 2011; Simmons, Daniel & Temple, 2019). As with each of these first four constructs, the Bayesian Network model revealed that the estimated probability of success increases when applied *in combination* with the other core constructs.

11.5.5. Coevolution constraints

Coevolution constraints was also hypothesised as a core construct in the CQIN model (i.e. it is a direct parent node of improvement outcomes). Unlike the preceding four constructs, this construct has a negative direction of influence. It was less strongly observed in the secondary data, indicated in 32% of cases. Potentially, negative factors are less likely to be described relative to the positive QI actions taken. The dynamic and subjective nature of agent network boundaries may also explain this finding. It is difficult to distinguish what constitutes growth or strengthening of one network from influencing an adjacent but separate network. Exploring this notion further requires more reports of unsuccessful QI where an inability to expand the agent networks is observed.

The deprescribing case study offered a clearer example of coevolution and the potential constraining effect on QI. An effective proof of concept for non-pharmacological support has been developed and tested. But scaling this solution any further requires substantial training and coaching of nurses and care home staff.

11.5.6. Problem space complexity

The term problem space complexity was chosen to reflect the idea that problem solving requires navigation of the many potential system states between the current state and the desired future state. This is based on Newell and Simon's description of 'problem spaces',

and how elements within these spaces are manipulated to solve problems (Newell & Simon, 1972). This is an appropriate proxy for complexity, as the problem space may be subject to any complexity conditions. The concept was not commonly referenced in the secondary case data, observed in just 34% of cases. This may be due in part to the somewhat academic framing of 'problem space'. To count in the thematic analysis, there needed to be some indication that teams were exploring the nature of the problems they were trying to solve and revisiting their initial assumptions. Another explanation may be the inverse relationship with problem solving effectiveness. Effective problem solving implies that problem space complexity has been successfully countered, even if not described in detail.

11.5.7. Learning Gain

Linked to all the constructs mentioned above is the concept of learning gain, identified in 34% of the secondary data cases. Learning is conventionally viewed as an output or product of quality improvement. Having participated in an improvement process, new knowledge is acquired and incorporated into current understanding of the process (Anderson, 1993; Lemieux-Charles et al., 2002). The CQIN model turns this into a continuous process focussed on the given problem. Whatever lessons are learned from one learning cycle are carried over into the next cycle. This process is of course no different to what Deming prescribes for PDSA cycles (Deming, 2000). The healthcare QI literature and the secondary case data both suggest that cultivating sufficient learning cycles is hard to achieve in practice (Reed & Card, 2016). However, without a mechanism to create and apply new knowledge, an unfamiliar problem space cannot be resolved. If the problem space is large, the time required to learn and resolve may be very long, and unrelated to any externally applied project timelines.

11.6. Theoretical Perspectives

This section now brings the results, themes, and literature together to place the research into context with established and emerging QI theory. The literature review and the secondary data analysis show that a broad and deep theoretical base exists for both QI and complexity. At the same time, there is a strong impression that these foundations are not yet widely operationalised into integrated frameworks for complex QI.

There were a small number of examples within the secondary data that used more integrated approaches when applying complexity principles to QI. Storkholm, Mazzocato

and Savage (2019) found that combining the sensemaking framework of Cynefin with Adaptive Leadership provided stronger coupling between problem and appropriate actions. This was a case of surveying the available tools and deciding to combine two of them. In other examples more novel approaches were applied, such as the application of the 'dissipative structures' model by Essén and Lindblad (2012), and an innovation diffusion model, based on virus spreading processes, proposed by Talbert-Slagle, Berg, and Bradley (2013). These two cases were examples of using general models (one from physics and one from biology) as *explanations* for an *ongoing process* of change in the healthcare QI context. As with CAS, the question of useful metaphor or literal explanation remains open, and the usefulness is the coherent set of ideas that comes with an integrated theoretical model, in contrast to focusing on isolated features of complexity. Uncritical translation of ideas across contexts is open to challenge; researchers and practitioners should also provide plausible explanatory mechanisms for the new context. This is what Booth, Zwar and Harris (2013) provide for their 'complex responsive process' of organisational change. In their model, change is a permanent, ongoing interaction between an environment and agents relating in communities within that environment. These examples of integrated complex QI thinking stand out amongst the tentative explorations and selective application of individual CAS principles that dominate the healthcare QI literature. After 20-plus years of healthcare complexity research, their rarity suggests that integrated modes of analysis remain underexplored.

11.6.1. A CAS Explanation of QI

Networks of adaptive agents receive, interpret, and share information within a given environment (their socio-technical context). Agents cooperate to achieve shared improvement goals. Variation, in the form of new information, is required to solve problems when the problem exceeds the current level of understanding within the network. New information must also be distributed sufficiently to reach a threshold where any new understanding can replace the status quo. An adequately replicated idea may still fail a selection 'test' within the operational environment. This may be because the new understanding is not valid (agents may be aligned but incorrect). Selection may also fail if the status quo remains better matched to the environment for other reasons. In other words, sub-optimal ideas may persist when the environment favours them. This description presents QI as a continual evolutionary process without end. Yesterday's successfully evolved improvement may fail tomorrow if environmental changes mean

that it is no longer fit in the sense of 'fit for purpose'. Evolution requires, at a minimum, variation, replication, and selection (Dawkins, 1976). Biological evolution utilises genetic information/encoding to transfer information. Information is also transferred culturally, and humans have accelerated this cultural transmission pathway via language and explainable knowledge (Dennett, 2017; Deutsch, 2011). However it is acquired, information accumulates into worldviews or mental representations – an agent's schema for interacting with the world (Dennett, 2017; Holland, 2012).

Agent networks do not exist in abstract space; they function within socio-technical structures. 'Structures' provide system order or regulation in the form of resources, boundaries, rules, and social norms for interaction. If socio-technical structures respond more slowly than the agent networks within them, high levels of complexity will stress the system regulatory mechanisms and inhibit the underlying evolutionary processes. Complexity conditions such as high rates of change, a large number of elements, a large number of relationships, interaction effects, and disorder, all affect the agent network itself (network stability) and the environment the network is operating within (external environment stability). In summary, complexity conditions can limit the ability of the agent network to exchange information, maintain shared understanding and cooperate within a socio-technical context.

A CAS explanation of QI suggests that high levels of complexity place pressure on the evolutionary capability of adaptive agents *and* the socio-technical context in which they operate. Protecting and nurturing the conditions for evolution (variation, replication, and selection) must therefore be the underlying principle for any CAS-informed QI methodology. The next two sections present the rationale for this assertion.

11.6.2. Premise 1: Change as an Evolutionary Process

From its earliest origins CAS thinking has provided holistic explanations for complex natural and social processes. CAS principles describe natural change processes, frequently in terms of ecosystems, but also referencing the underlying evolutionary processes (Holland, 2006; Kauffman, 2000). Observation and analysis at finer granularity reveals the biological, chemical or physical foundations (Krakauer, 2023). Nevertheless, the commonly described CAS characteristics function predominantly within an ecosystem framing and are accessible to analysis at this level: Local patterns of behaviour; interactions between autonomous agents following simple rules; and self-organisation

into small groups. Feedback effects and instability in the environment then lead to a continual process of change and adaptation. This is uncontroversial when describing a natural ecological system. What constitutes this evolutionary process in our economies, cities, and organisations? The transmission of *information* is the common link between biological and social evolution (Adami, 2012).

Cultural transmission gives humans the ability to share information about our environments in an alternative pathway to genetic encoding. Storing information then enables learning beyond direct observation and imitation. This cultural evolution augments biological evolution and has enabled human flourishing. Cultural evolution is an established concept, but the consequent idea that our social structures might need to protect and facilitate this process is encountered less frequently (Creanza et al., 2017; Dawkins, 1976; Miton & DeDeo, 2022; Smolla et al., 2021). Knowledge is a cultural artefact, and the evolution of knowledge relies on the same evolutionary drivers of variation, replication, and selection. Changes in the environment generate variation. Replication of information is not generated by the information itself, but by the networks of agents sharing and updating their collective world view or schema (Dennett, 2017; Gabora, 2004). The updated schema may or may not flourish; it is continually subject to tests of fitness within its context. In this manner, improved ways of working gradually supplant earlier ideas, to be replaced themselves by the next generation of knowledge. The transition in knowledge is not inevitable; inferior ideas may remain better adapted to the environment and persist (Smerdon et al., 2016). Any objective of long-term, sustained improvement is also contingent on the stability of the environment.

Within the CAS explanation of QI, the broader concepts of culture and information are narrowed to collaborative problem solving amongst a group of agents. Despite popular conceptions of evolution as a competitive process between individuals, cooperation is an effective evolutionary strategy and becomes necessary to build order amongst individuals and groups of individuals (Nowak, 2006). Cooperation also becomes self-reinforcing when the shared benefits increase over time and groups become more skilled at cooperation (i.e. problem solving). Nowak (2006) outlines five mechanisms for cooperation that move from the very local to the more distant: (i) kin selection, (ii) direct reciprocity, (iii) indirect reciprocity, (iv) network reciprocity and (v) group selection. Cooperation mechanisms take place on a continuum that leads to increasing scale and more sophisticated order within groups (Dennett, 2017; Nowak, 2006; Wilson & Kirman,

2016). Therefore, cooperation is a necessary *scaling* process for the emergence of system-level change.

The idea that this fundamental evolutionary process breaks down with the introduction of human self-awareness and agency can only be supported if an argument for human exceptionalism is accepted. This was the initial objection to healthcare as CAS raised by Paley and Eva (2011). The counterfactual argument is that agency exists at multiple levels of sophistication, and goal awareness simply increases the feedback and interaction effects within our complex social systems (Dennett, 2017; Tomasello, 2022; Wilson & Kirman, 2016). The intentionality of a QI endeavour is not a forbidden example of design within a CAS. Rather, it is an example of self-organisation amongst adaptive agents creating a local optimum state.

11.6.3. Premise 2: Socio-technical Structure as Inhibiting or Enabling CAS

The social CAS outlined thus far requires an environment for the agent networks to function within. Socio-technical structures provide the resources, boundaries, and ordering rules for the agents. As the ordering structures grow, their complexity also increases. Expansion of variety, leading to complexity, occurs inevitably as a function of increasing system components (Ashby, 1956; Kauffman, 2000). Therefore large socio-technical systems must manage increasing variation. The problem is long recognised within organisations; viable systems modelling, variety engineering and self-organising teams are examples of responses to this phenomenon (Ashby, 1956; Leonard, 2009; Trist, 1981). Regardless of the strategy taken to achieve regulation, Ashby's concept of requisite variety is key. If the socio-technical regulation cannot respond to the variety it is subjected to, it will become sub-optimal. High complexity, demanding high variety management from the socio-technical system, is likely to overwhelm the system regulatory functions, leading to undesirable outcomes and barriers to change. Borrowing the language of Holland's (2012) concise summation of CAS as *signals and boundaries*, system signals will be degraded and boundaries unable to be navigated.

The need to manage organisational barriers helps to explain the development of complexity-aware improvement frameworks focusing on leadership. Yes, as Carmichael & Hadžikadić (2019) propose, introducing leadership into CAS may just be an inherent bias towards hierarchy and unnecessary complication. But an alternative interpretation is that focusing on leadership strategies and actions is a very realistic response to the power

dynamics within modern organisations. For better or worse, hierarchical leadership is often necessary to change organisational rules. In the words of one recent review into new Zealand's health reforms:

The role of local and central agencies and the government is to provide the policy settings and conditions in which trust-based networks can flourish. (Sharma et al., 2024 p. 11).

Socio-technical theory also recognises the value of flexible working group configuration and at least some level of decision making delegated to small operational groups (Appelbaum, 1997; Trist, 1981). In other words, self-organisation is an effective evolutionary strategy, allowing the flexibility and cooperation necessary to respond to changes that cannot be predetermined. Change led by senior management and Kaizen-style small group change each have a place in this ecosystem.

11.6.4. Assessing CQIN as a CAS Theory of QI

The theory underpinning CQIN is expressed as Hypothesis 4:

(H4) *Within socio-technical CAS, system-level change in the form of goal-directed improvement is achieved through successfully coordinated updates and alignment of agent schema.*

Definitions of the constructs and boundaries for performance have been established. The prediction within H4 has been supported by the simulation experiments and the secondary data analysis. The basic requirements for a useful theory have been met (Wacker, 2008). A description of *how* H4 works is now required.

The adaptation of a complex adaptive system involves learning (Doolittle 2014; Holland 2006, 2012). If a chosen improvement action is working, any progress is then understood via feedback signals and incorporated into subsequent actions. If the feedback paths are not in place, or the signals invalid or lost in noise, the information required for updating the agent network understanding is compromised. If the chosen action is not effective, either the implementation or the underlying change hypothesis needs to be revisited. *Double loop* learning is not new idea (Ashby 1956; Argyris 1976; Umpleby 2008), but seemingly hard to achieve in QI practice (Reed & Card 2016). The requirement to continually revisit the goals and actions may sit at the heart of the critique of conventional QI methods in complex scenarios. Regardless of the conceptual strategy or

method applied, an improvement team genuinely reviewing and updating their efforts is likely to have more success than a team focussed on one solution compromised by rapidly changing circumstances or incomplete information.

In this iterative manner, an agent network can progress through the available problem space until an effective solution is found (Anderson 2005; Newell & Simon 1972). Provided there is sufficient time and resources available, even a low-skilled agent network could theoretically make some progress towards their improvement objective. However, there are very real limits on time and resources. Increased time exposes the CAS to the risk of change in external conditions. Equally important, the ability of the agent network to maintain focus and a shared schema on any one QI objective may also diminish over time. These challenges can be mitigated by deliberately building a genuinely shared schema amongst agents and using structured problem-solving methods, for example, rapid hypothesis testing iterations, to generate and verify new knowledge. To be effective in overcoming complex problems beyond the current state of knowledge within the agent network, practitioners must ensure that newly acquired knowledge is not lost or isolated. Appropriately integrated into an agent's schema, new knowledge informs the next learning cycle, creating a coherent, gradually compounding force for change. (Dixon-Woods & Martin, 2016; Kovach & Fredendall, 2013; Kringos et al., 2015). Without an *amplifying* mechanism, the gap between current and target future system states cannot be closed.

Collaboration, problem solving, and effective communication are all essential for effective QI. This is of course a very orthodox description of the improvement process. CQIN theory posits that a shared understanding of the problem is the *sine qua non* for the large-scale cooperation required across large agent networks. Maintaining and growing the shared understanding — the 'what, why, and how' of a given QI problem — focuses the above actions on closing the representation gaps within agents' shared schema (Cronin & Weingart, 2007; Choi et al., 2001; Doolittle, 2014; Nordin, 2017).

CQIN is literally an evolutionary process. *Variety* must be present in the form of new information, *replication* takes place as the minimum level of shared understanding is grown across the agent network, and *selection* occurs only if the changes are effective, displacing the status quo and functioning as intended. Unintended consequences may emerge, as well as new changes in the environmental context. Therefore the learning

process is never ending, and critically, it is embedded within the overall activity of the agent networks. The very tight coupling between ‘doing’ and ‘improving’ explains the difficulties of ‘imposing’ change from outside, and the ongoing frustration with management structures that see this as the optimal approach (Braithwaite, 2018).

As an evolutionary strategy, continual learning cycles maintain and strengthen an agent network’s ability to evolve in response to environmental signals. The actions of cooperation and problem solving build teamwork and the conditions for shared understanding. In simple terms, alongside any given QI goals, cooperation itself improves with practice (Coleman & Wiggins, 2017; Lemieux-Charles et al., 2002; Nordin, 2017).

Figure 11.1 revisits the original CQIN conceptual model (Figure 4.1) now as an evolutionary process. Figure 11.1 shows the relationships between the background variation, replication, and selection evolutionary phases alongside the information exchange, cooperation, and learning activities taking place. This macro-level (system level) view of CQIN complements the micro-level provided at the agent level. Both representations should be thought of as one still ‘frame’ or moment in time within a dynamic, continuous process.

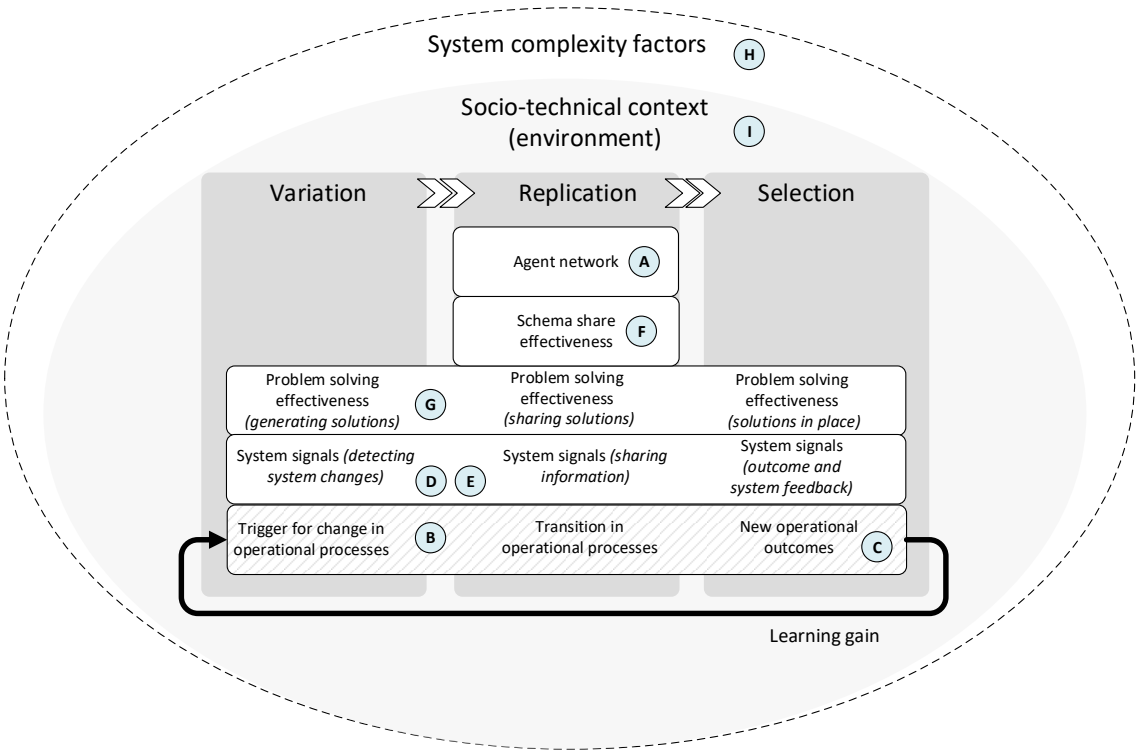


Figure 11.1: CQIN as an Evolutionary Process

System signals and problem solving have different functions within the evolution phases. In the variation phase, new system signals are detected. These signals may trigger a need or perceived need for change, which requires solution generation. In the replication phase, the solution options are shared more widely amongst the agent network until the threshold for change, unique to this context, is reached. In the selection phase, the new operational state must demonstrate it is fit for purpose in its operational context, with maximum wanted outcomes and minimum unwanted outcomes. Unless the problem being worked on is trivial, the progress made and any lessons learned need to become the point of departure for the next change iteration in an ongoing process of refinement. The socio-technical context and system complexity factors moderate these learning cycles, enabling or inhibiting the ability for relevant system signals to circulate and for agents to cooperate.

11.7. Implications

Having set out the results of this research within a theoretical framework, it is timely to revisit key themes within the healthcare QI literature to see where the results and proposed explanations are consistent with this literature. This study supports current discourse, but also introduces some new questions. It is proposed that there are several theoretical implications that arise when QI is interpreted as a CAS. 'Interpretation' is used here to mean moving from using CAS as an analogy (QI functions like a CAS) to using CAS as an explanation (QI is a CAS).

Overall, this research lends support to the assertion that healthcare is a CAS, by finding empirical support for a socio-technical CAS explanatory mechanism. No strong counterfactual causal explanations were revealed, but the answers to questions over the suitability and effectiveness of QI methods are more nuanced. The scoping review and secondary data analysis revealed that descriptions of 'complexity' are vague, inconsistent, and possibly overstated. Accepting some level of complexity to be present, skill and fidelity to improvement method principles clearly continue to have a positive influence on QI success. Conventional QI remain methods remain highly relevant.

One unanticipated finding in the secondary data was the direct importance of skills development, covering both QI skills and professional clinical development. Professional skills development and focused practice increase the exposure to the problem spaces likely to be encountered, and the available resolution options. This is *deliberate practice*

and *skills-based quality* in action (Beckford, 2002; Ericsson, 2004). Collaborative learning amongst professional groups is a legitimate response to complex healthcare quality improvement (Haraldseid-Driftland et al., 2022).

An underlying but neglected evolutionary process may also help explain the apparent lack of 'transformational' change in healthcare quality in comparison with other industry sectors (Dixon Woods & Martin, 2016; Mondoux & Shojania, 2019). The necessary timescales may not be appreciated, and the immense range of healthcare QI activity taking place must be bound together in some manner to build transformational change. Without the ability to learn from all this activity, it remains a collection of local optima and not a system-level optimisation. The need to coordinate and deliberately scaffold towards new solutions — a learning gain — is recognised in the literature (Dixon Woods & Martin, 2016).

Proponents of CAS theory assert that elements of the system can be manipulated to achieve desired change, despite the difficulty of identifying appropriate interventions (Britto et al., 2018b; Khan et al., 2018; Holland, 2006, 2012). Success in this context means adaptation and fitness for the environment. In addition to the challenge of sensemaking under conditions of complexity, the key difficulty left unexplained is identifying what the leverage points are. CQIN, as a CAS theory of QI, suggests that the starting point for system leverage is the underlying evolutionary process. Where and how are variation, replication and selection being inhibited? In the QI domain, this means observing, understanding and (if required) manipulating the *schema share effectiveness, agent network communication paths, problem solving effectiveness* and *system signals.*

Positing an underlying evolutionary process also leads to specific implications not commonly discussed in the healthcare QI literature. Five of these implications are now described in more detail.

11.7.1. The Gap between Process and Improvement

An evolutionary process challenges the notion of a gap between process and improvement. Can working 'in' a process and working 'on' a process be meaningfully decoupled? Accepting that this decoupling is difficult may explain why EBM is so difficult to achieve in practice. At the same time, the championing of self-organisation by healthcare as CAS proponents gains more plausibility if it is only the agents involved who can ever really optimise a process by evolving to new states. The need to take on new

knowledge and apply it remains, but in the evolutionary view the external ideas and expanded agent networks are simply supplying variation and additional resources for problem solving and cooperation within the network. This framing is consistent with Hawe's (2015) emphasis on incorporating field observation and shifting of practitioner mental models in complex interventions. Further evidence that the gap between doing and improving is illusory was suggested in the secondary case studies with *workflow integration*, *self-directed learning* and *professional/capability development* identified as additional themes for QI success.

11.7.2. Path Dependence

Explicit references to path dependence are much less common in the healthcare as CAS literature than the descriptions of self-organisation, non-linearity, or emergence. Yet path dependence has profound implications for any QI that does not acknowledge it. Expression of a target or goal state is almost invariably the starting point for a QI solution. When the gap between current and goal state is conceived as a problem space to successfully navigate, emphasis returns to the *starting point* for change. A path between states may have to be built and require multiple iterations and development of fundamental prerequisites. In CAS theory, not all possible states are even accessible, Kauffman's memorable phrase "the adjacent possible" is optimistic (change is possible) and at the same time states a clear rule (there must be a path to the new state, one step at a time (Kauffman, 2000). Therefore, A CAS interpretation of QI assumes moving from current to goal system states is highly conditional and takes place as a sequence of discrete learning steps. Adequate resources may accelerate the process, but the underlying adaptive mechanism does not change.

11.7.3. The Role of Organisational Structure

What function do conventional socio-technical structures and methods have in a CAS interpretation of QI? A CAS explanation of change suggests that if fundamental requirements for evolution are inhibited by the socio-technical context, the likelihood of quality improvement success is reduced.

'Organisational structure' is used here to refer to how organisations are set up to function. This may include hierarchical leadership layers, structural grouping of teams, codified rules and norms, and methods for communicating and managing change. This study did not examine whether established organisational structures, policies and change

strategies are necessary or sufficient in their own right. However, cooperating networks of agents must function within and across these organisational structures. The structures provide an environmental context for the CAS, using the CAS_2 'network of adaptive agents' definition of Wilson and Kirman (2016).

In the organisational context, rules coordinate behaviour, and these are necessary constraints for a minimum level of order (Rancati & Snowden, 2021). However, the organisational structures and methods will further constrain information flow and cooperation in the presence of complexity. This is because the agent network and organisation respond to change on different time scales. A team within a clinical microsystem may quickly adopt a new process, but necessary changes in funding, regulatory policy or supporting technology are unlikely to work at the same speed.

An underlying assumption should be surfaced at this point in the discussion. Implicit in the consideration of organisational context is the belief that organisations are rational and that they do indeed seek to improve. Selznick (1948) argues that rationality cannot be taken for granted and that organisations are inherently dysfunctional, resulting in irrational behaviour and sub-optimal outcomes. It should not be presumed that organisations possess the motivation to improve (Selznick, 1948). The argument applies to QI in general and has not been critically examined in this research, which follows the domain orthodoxy in believing that individuals, and collections of individuals, do seek purposeful change for the better within their organisations.

A positive role for appropriate leadership within CAS is finally apparent. Unlike natural CAS, Socio-technical CAS require leadership to effect any necessary changes within organisational structures. Complexity frameworks emphasising leadership actions fulfil an important role. Rather than being a misreading of essential CAS principles, the requirement for leadership arises from the need to modify the environment the CAS inhabits. The relative scarcity within healthcare of enduring examples of the viable system model, or of Senge's (2006) learning organisation, also suggest that deep structural change of organisations at this scale is difficult to achieve in practice.

When there is a mismatch in the rate of change between organisation and agent network, two general strategies seem possible. One approach is to increase control and formalisation via structure and hierarchy. This may take the form of minimising variation and suppressing change. Where changes are seen as necessary, system regulatory

mechanisms must be updated to match the new, required variation. This is organisational cybernetics and the viable systems approach to complexity. At the risk of stating the obvious, complexity increases as a function of the number of variables and potential system states (Ashby, 1956; Kauffman, 2000). In this study, the three most common conditions for complexity identified in the secondary case data were high numbers of system elements, high numbers of interactions, and high diversity of elements. All these factors increase variation. In a cybernetic system, the ability to regulate high variety systems is necessary and few healthcare organisations are set up to achieve this (Akmal et al., 2021; Leonard, 2009).

In contrast to attempting to minimise and/or regulate variation, a second strategy for managing the organisational constraints arising from complexity is to accept that a level of uncertainty will always be present and to proceed pragmatically from this understanding. Organisations, teams, and individuals need to be able to function with incomplete information, make decisions, and update their understanding when more information is available. In terms of CAS principles, self-organisation is natural in this context. Acknowledging this uncertainty is explicit in the Cynefin framework, where the management response strategies depend on the level of knowledge. There are two complementary, non-CAS supporting arguments for this second approach. Heuristic decision making recognises that information will always be incomplete (bounded rationality). At times, decisions must be made with incomplete information, and with the right heuristics this can be as successful as methods utilising more data (Gigerenzer & Goldstein,1996; Simon, 2000). Similarly, Bayesian reasoning acknowledges that subjective judgement is necessary, and that understanding is built up iteratively as more knowledge is gained (Pearl & Mackenzie, 2018).

If the critical role of self-organisation under conditions of complexity is accepted, it is appropriate to ask to what extent organisational structures can be re-arranged to facilitate this self-organisation. Organisational designs built to enable collaboration amongst autonomous agents have been proposed, such as *actor-oriented architecture* (Fjeldstad, et al., 2012). In this approach, a knowledge ‘commons’ is created, along with necessary policies and processes, to support the circulation of necessary information (system signals) and self-organisation amongst actors. The Viable Systems Model is also an example of a socio-technical structural design intended to support continual change and self-organisation (Akmal et al., 2021; Beer,1984; Leonard, 2009). The question

remains open whether new organisational designs can meaningfully reduce unintended constraints to CAS functioning.

11.7.4. Wicked Problems

The CQIN model posits that a minimum level of shared understanding across large, heterogenous networks is a prerequisite to cooperation. This proposition was supported in the secondary data by the observed coverage of the *schema share effectiveness* construct. It introduces a sobering note of realism to 'wicked' problems. Shared understanding, to a sufficient level to achieve cooperation, may simply be impossible in some circumstances. In the words of Rittel and Weber, planning should be

> *…an argumentative process in the course of which an image of the problem and of the solution emerges gradually among the participants, as a product of incessant judgement, subjected to critical argument* (Rittel & Webber, 1973, p. 162).

This process, if successful, 'tames' the problem, allowing problem solving to begin. The high-level process described is very similar to soft systems methodology (Augustsson et al., 2020; Checkland, 2000), and the 'complex process of relating' described by Booth et al., (2013). Healthcare complexity researchers are aware of the difficulties wicked problems pose and have attempted to resolve these with appropriate frameworks. Despites these attempts, some wicked problems cannot be tamed (Rittel & Webber, 1973); CQIN supports this assertion by proposing that shared understanding must reach a minimum level for system change to be viable.

11.7.5. QI Success Understood as a Probability Distribution

Following on from the acceptance that not all problems can be solved is the idea of probability of QI success. Improvement practitioners tacitly understand project success is not 100% certain. Between the 100% level of guaranteed success and the 0% level of a wicked problem unable to be tamed, there is a Probability Distribution for QI success. This not an intuitive notion – how is the probability of a unique event determined? The conventional definition of *relative frequency of occurrence within a reference class* (Gigerenzer, 1994) applies but also seems inadequate, as it expresses probability at the population level across a group of similar events, and not a for a single QI project. Nevertheless, this is exactly how most healthcare risk is calculated.

Thinking probabilistically about QI success recognises that the probability of success cannot be 100%. From that start point Bayesian reasoning can be introduced to focus on the unique instance being engaged with. Where does this individual project sit within a reference class of comparable examples? What local expertise and knowledge do we have that can reliably establish the probability? What steps can then be taken to improve the probability of success *in this instance?* This research has demonstrated that a reference class of complex projects can be defined and used to support Bayesian reasoning.

11.8. Opportunities: Practical Applications

Alongside the theoretical considerations, a CAS explanation of QI generates practical opportunities for improvement practitioners. Five broad opportunities/potential applications are set out below.

11.8.1. Evolutionary Foundations Support Diverse QI Methods

The evolutionary principles of CAS are not in conflict with improvement methods and tools. Technical interventions are still going to be required. The important point is that alongside any conventional interventions, the conditions for networks of adaptive agents to flourish must also be in place. Targeting these evolutionary foundations becomes a legitimate QI approach. Examples of this were seen within the secondary case data themes. Improving the coordination/navigation amongst different professional groups was seen as a practical QI intervention in its own right (Altman, Woolfenden & Breen, 2017b; Bishop & Waring, 2019; Praetorius et al., 2018). A separate example is the long-term nurturing function of capability and skills development (Goderis et al., 2009; Pesec et al., 2021; Riordan et al., 2019; Quinn et al., 2021).

In CAS terms, these strategies improve agents' abilities to identify problems and resolve them cooperatively. Focusing on the number, type, and strength of connections across agent networks becomes a practical QI tool. Similarly, ensuring healthcare professionals have the cognitive space and learning environment to practice problem solving within their regular professional activity.

11.8.2. Value of Existing QI Methods and Thinking

The value of both traditional and complexity-focussed QI methods and frameworks was strongly supported by this study. Across the many geographical regions and healthcare improvement contexts, successful application of very conventional QI methods such as PDSA cycles or Lean management was widely reported. The difficulty in establishing if

these really were complex scenarios is acknowledged, but the consistent level of observed use demands due consideration. Equally important, the wide range of methods reported demonstrates that there are many potentially successful QI tools available and in use for nominally complex situations (Figure 3.3, Section 3.6). It is possible to view this range of QI methods as an ecosystem of approaches relevant to many different contexts, but with considerable overlap of underlying principles.

Conventional QI tools are practical problem-solving methods. Application into complex scenarios requires additional techniques, but this is an expansion of approach rather than rejection. QI teams are using complexity-informed approaches, situational awareness and sense-making as a diagnostic front end before selecting a course of action (Storkholm et al., 2019). PDSA cycles can be used to evaluate a problem space and generate solutions (Czulada, Leino, & Willis, 2016; Hoyle, Patino, & Yoxall, 2019; Pearson et al., 2016). As a technique for QI, Kaizen is well matched to the complexity principle of self-organisation. Again there is evidence of its effective use in healthcare QI (Erdek & Pronovost, 2004; Ferreira, & Saurin, 2019; Radcliffe et al., 2020). The secondary case data suggests that QI practitioners have pragmatically combined proven and familiar tools with complexity-informed thinking.

There were multiple examples of complexity-informed leadership being applied (Hilts et al., 2012; Shelby et al., 2019; Vatnøy, Karlsen, & Dale, 2019). These examples show that practitioners are already working with these ideas. Paying due attention to appropriate leadership remains an important practical QI element. Again, this is hardly a new idea. However, the type of leadership and leadership actions in the context of complexity are key. Viable complexity leadership frameworks exist and their use can be expected to increase. The organisational context is the environment that influences the agent networks — creating the conditions for change to flourish is a leadership function.

11.8.3. Focus Areas for QI Activity

Alongside a confirmation that conventional QI methods still have a place in complexity-informed strategies, additional focus areas were also suggested by this study. Professional development and self-directed learning have already been noted. Also apparent within the secondary data cases was the role of *system signal timing and integration* (Alhaider et al., 2020; Elertson, 2017; Murray et al., 2011; Simmons, Daniel & Temple, 2019). These examples demonstrated that information can be an improvement lever in its own right,

provided it is timely and properly integrated into process workflow. In other words, ensuring staff have the signals to trigger and guide action when it is needed. Many organisations might assume this would already be the case; examples of signal integration as a practical QI strategy suggest otherwise. These three concepts are not new but the reminder of their value is important.

The idea of phase transition (non-linearity) between changes in system states is ubiquitous in the complex system literature. The simulation experiments demonstrated how these step changes can occur in a socio-technical CAS. Embracing non-linearity in change progress also presents as an important practical QI focus. If thresholds for system change exist, then identifying progress in relation to these thresholds is critical. QI teams can explore where their unique system transition points are and take actions to reach and maintain minimum levels of the relevant CQIN constructs. An understanding of non-linearity should be incorporated into change project reporting and management. Progress and risk assessment for genuinely complex scenarios are incomplete without this understanding.

11.8.4. Using a BN model of QI

The CQIN BN model is a practical risk analysis tool for complex QI projects. The representation of a system as a BN enables Bayesian reasoning to update understanding as more information is obtained. Network fragments can be manipulated to better reflect direct and indirect influence amongst system variables, and individual nodes deconstructed to the level of granularity required. For example, if a QI team judges that its level of shared understanding (schema) needs to be improved, contributory variables like *interpersonal factors* can be broken down further and explored. Are practical steps to increase social interactions and trust required? Intervention strategies can be modelled without disruption to real-world activities. Gathering participants to assess and apply values to the constructs is a sense-making and situational awareness activity exactly as for a method like Cynefin. Discussion and debate will follow, leading to better awareness of the representation gaps within the group (Cronin & Weingart, 2007). Are our *agent network communication paths* enabling or inhibiting the *system signals*? What practical steps can we now take to increase the node degree across the network?

11.8.5. Practical Descriptions of Complexity

The value of Wiesner and Ladyman's (2019) framing of complexity measures into 'conditions for complexity' and 'products of complexity' was constant across the different stages of this research. It is certainly possible for QI researchers and practitioners to build up a picture of complexity by engaging with the theoretical and applied complexity literature. However, this is a slow process, and as observed in the secondary data, much of the complexity language within the healthcare literature is inadequately defined. The approach taken by Wiesner and Ladyman identifies conditions that cause complexity, along with characteristics of complexity that may be observed. Their work is relatively recent and fully consistent with the wider complexity science cannon. It represents a step forward for practical complexity research. Are characteristics like self-organisation and non-linearity observed? Is this being caused by high numbers and diversity of system elements, feedback effects or a lack of ordering constraints? Such distinctions make it easier for practitioners to identify complex systems and potential improvement options.

11.9. Summary of Discussion

This chapter has covered many discussion items. The notable research findings of: (i) non-linearity; (ii) viability thresholds; (iii) learning gain; (iv) sets of sufficient conditions for success; (v) directions of influence, and (vi) mediation and moderation effects were introduced first. These findings feature throughout the discussion. The consideration of the research hypotheses and the research question followed. The research hypotheses were deemed to be supported, and the research question answered via the results for the individual CQIN constructs.

A theoretical perspective was set out, where the CAS explanation of QI was presented as an evolutionary process. Foundations for this interpretation and how it works as an explanation (not a metaphor) for socio-technical contexts were detailed. Finally, implications for the theoretical discourse and practical opportunities were proposed. Considerations for current healthcare QI discourse are: (i) the coupling between process and process improvement; (ii) what path dependence really means; (iii) the role of organisational structure; (iv) wicked problems, and (v) the notion of QI success as a probability distribution. Practical opportunities proposed were: (i) the nurturing of evolutionary conditions alongside existing methods and frameworks; (ii) confirmation from the applied literature that existing methods have a role in complex QI; (iii) novel

focus areas, especially non-linearity and signal integration; (iv) using a BN model as a sense-making and risk assessment tool, and (v) distinguishing between the causes and characteristics of complexity.

Chapter 12 Conclusions

12.1. Introduction

This research has explored the potential of a CAS interpretation of QI. An explanatory model for *how* QI functions as a CAS was developed and tested, the *complex quality improvement network* (CQIN). Applied healthcare QI case studies were examined for evidence of the proposed CQIN constructs. The CQIN model is undoubtedly rudimentary but it is consistent with the empirical data. CQIN was conceptualised to explore themes within the healthcare as CAS literature. It is encouraging to find experimental and empirical support for those themes, and to add the perspective of this research to a wider conversation.

Foremost amongst the themes discussed in the previous chapter is the utility of *integrated* theories of complexity-informed QI. CQIN presents a theory of QI built on the evolutionary principles of CAS, positioned within human socio-technical systems. A second essential thread in the discussion is the role of traditional quality management methods and tools in situations of high complexity. This study has found that healthcare QI practitioners are routinely expanding their toolkit to respond to complexity and that conventional QI techniques remain essential. However, these techniques must be augmented with approaches that facilitate sensemaking and a realistic and humble mindset when working with incomplete information.

As an integrated theory of change, CQIN highlights a potentially underused strategy for healthcare QI practitioners. This theory proposes that nurturing the evolutionary conditions for change must take place alongside any other change endeavours. Building an aligned understanding amongst all participants is a prerequisite for the large-scale cooperation necessary to deliver (and improve) healthcare services. Communication pathways must have sufficient capacity for the rapid sharing of system feedback and updating of collective schema. Any new system state is not a destination, but a transient step in an ongoing process of adaptation within a constantly changing environment.

12.2. Review Against Research Objectives

Five research objectives were set out at the beginning of this study:

1. To investigate the role of system factors as possible components of change within CAS.

2. To examine the interaction of system factors within CAS.
3. To explore the relative influence of system factors as components of change within CAS.
4. To investigate factors that may constrain change within CAS.
5. To validate and verify the results using simulation and secondary data from published case studies in healthcare QI.

To what extent were these research objectives achieved?

Objectives 1 and 2. Components of change and system interactions within CAS

The CQIN model identifies a minimum set of factors for a sociotechnical CAS, based on a network of adaptive agents engaged in cooperative problem-solving. Agents require communication paths (a network) to share information and cooperate. Feedback signals must flow around this network to allow agents to update their schema. A threshold of shared problem representation must be reached for cooperation – the *schema share effectiveness*. Problem-solving actions are generated from the current agent schema; if insufficient to resolve the problems being worked on then new knowledge must be introduced. This is either generated through learning activity or imported into the agent network.

Objective 3. Relative influence

The previous paragraph describes four core CQIN constructs: (i) schema share effectiveness; (ii) agent network communication paths; (iii) problem-solving-effectiveness, and (iv) system signals. These were the top four factors identified in the secondary data analysis, all featuring in greater than 50% of reviewed secondary cases. The level of abstraction is deliberate and each of these factors can be broken down into contributing factors. For example, *problem solving effectiveness* requires inter alia, focus, autonomy, existing knowledge, the ability to generate solution options, and the ability to learn. The granularity of the factors offers an insight into the overall influence. For example, interpersonal factors are posited as an important contributing factor to the schema share effectiveness. Improvement teams who identify issues of trust or psychological safety amongst a group of agents will reasonably seek to address these in the expectation that this will in turn lead to improvement outcomes. The CQIN framing reveals the limited marginal influence of any one factor. Improved trust amongst agents

might allow the schema share effectiveness to reach a threshold for change. However, if the given scenario is also dependent on better system signals and new knowledge, the probability of success is only marginally increased.

The QCA revealed that each of the four core CQIN constructs appeared to be the influential success factor in different cases. This finding can be interpreted in two ways. Potentially, the other factors were also present, at sufficient levels for success, but not observed or described. Alternatively, the problem scenarios could be thought of as systems in a semi-stable state where one construct has fallen below threshold. In these cases, paying attention to one construct was all that was required. Whichever interpretation is correct, assessing and optimising these four core CQIN constructs is a key strategy for healthcare QI practitioners. The Bayesian network analysis showed that the probability of QI success increases with each additional construct being observed as 'True' (i.e. above threshold).

Objective 4. Constraining factors

There are two further CQIN constructs hypothesised as core, *coevolution constraints* and *problem space complexity*. The empirical evidence for these constructs in the secondary data was considerably less than the other core constructs. As noted, each of these inhibiting constructs exists as a complement to a positive construct. Effective problem resolution implies that a given problem space has been explored and resolved, even if this process is not always described. Similarly, the idea of a coevolution constraint is complicated by the subjective interpretation of network boundaries. Expanding an agent network to include a new professional group that was inhibiting change may well resolve this constraint.

Central to this study has been the positioning of CAS theory in a socio-technical environment. The ability of socio-technical structures to respond to complexity presents perhaps *the* critical constraining factor to complex QI. Organisational inability to adapt quickly and continually to change explains to some extent the popularity of complexity leadership frameworks, as leadership is the conventional starting point for organisational change.

Objective 5. Validation

Validation of the CQIN theory of QI is supported within the boundaries of the study design and acknowledged limitations (12.4). Each of the four research components (simulation modelling, secondary data analysis, applied case study, and Bayesian network analysis) offered perspective and detail within a consistent theoretical framework without contradictory findings. The conceptual model offers a view of healthcare QI that appears well supported in the relevant case study literature.

12.3. Contributions of this Research

This study was motivated by the seemingly slow progress of healthcare QI in lifting health outcomes, coupled with a clear line of enquiry to pursue — providing more detail on how complexity-aware improvement methods might function. How do the findings contribute towards these aims?

12.3.1. Contributions to Academic Discourse

Quality management thinking remains a core component of complexity-informed improvement. Additional strategies are required alongside conventional QI methods and tools, especially sense making and complexity leadership. The utility of existing complexity frameworks is supported by this research and this is not surprising. Looked at closely, frameworks such as *soft systems methodology*, the *system of profound knowledge*, and *Cynefin* each contain this mix of sensemaking strategies and improvement tools. This study suggests the academic conversation should move beyond 'either/or' descriptions of conventional QI methods vs complexity frameworks. Within the specific boundaries explored by this study, the strong suggestion is that conventional QI remains highly useful when applied in the context of a CAS framework.

CQIN identifies an adaptive process applicable to socio-technical CAS – problem-solving within networks of cooperating agents. Without such an explanation, claims that socio-technical systems are literally CAS are unsupported. New improvement strategies become available if our understanding eventually moves from healthcare QI functions *like a* CAS to healthcare QI *is a* CAS. This research is insufficient on its own to prove that case, but it has found theoretical and empirical support for the literal interpretation.

The frameworks noted above tend to respond to complexity through socio-technical filters. In other words, there is an emphasis on organisations, structure and the role of

leadership. CQIN is a framework that does not have this perspective. Yet clearly leadership and organisational structure are essential to socio-technical CAS. CQIN proposes a new relationship, where the socio-technical context is part of the external environment of the agent network. This environment may constrain the agent network because it cannot respond quickly enough in complex conditions. Looked at in a more positive light, the socio-technical context, together with complexity-informed leadership, might also *enable* adaptive networks. Organisational frameworks supporting self-organisation, such as the *viable systems model* or *actor-oriented architecture*, still have an important role to play in supporting CAS.

CQIN bridges a granularity gap in the healthcare QI literature, between high-level discussion papers advocating for complexity-aware approaches as one level, and detailed case studies as another level. The difficulty for practitioners is moving from the high-level statements around sensemaking, for example, and then finding practical steps to take. Adopting an existing 'off the shelf' complexity management framework is one option. This research used CAS theory to open up other strategies. Within the sets of sufficient conditions identified in the secondary data analysis, tangible improvement options are presented: improving combinations of *shared schema effectiveness, agent network communication paths, problem solving effectiveness,* or *system signal integration* will lead to improved QI outcomes. These constructs provide focus areas to start building more detailed interventions.

Another academic contribution is made in the area of reporting QI and complexity details within QI projects. This study found a low level of standardisation in complexity terminology and the reporting of key QI contextual features, consistent with previous reviews. For healthcare QI researchers, the CQIN model could be a unifying lens when applied to the healthcare QI literature, drawing together common threads across the long tail of diverse healthcare improvement settings and methods. This lens has value for future systematic reviews. In complex systems, intervention and context cannot be easily separated, and every case is unique. A coherent arrangement of fundamental principles helps categorise distinct and irreducible examples, using the CQIN constructs. The usefulness of Wiesner and Ladyman's (2019) approach to measuring complexity is also acknowledged. Distinguishing between conditions and characteristics of complexity is noted here again as it was an important enabler of the research, charting a path through the complexity terminology maze.

A final contribution to theory is the potential generalisability of CQIN to knowledge domains beyond healthcare QI. Complexity is certainly not limited to healthcare. Healthcare QI was the focus area and proving ground for CQIN, but the model, as with CAS theory, is a general model, and could easily be reframed in other socio-technical contexts such as social policy development or public service organisations. The need to successfully coordinate large numbers of diverse agents in the presence of high rates of change, conflicting goals, multiple feedback paths, and poor information flows, is a recurring feature of large-scale change endeavours.

12.3.2. Contributions to Healthcare QI Practice

The concept of sensemaking is perhaps the most prominent feature of complexity-informed improvement methods. CQIN is also a sensemaking tool and CQIN provides a practical structure for system observation and analysis relevant to the underpinning requirements for cooperative problem solving. The major practical contribution of this research is a diagnostic and predictive model for complex QI. CQIN shows plausible utility, within the boundaries of this study (simulation experiments and a sample of international healthcare QI projects, with self-reported complexity factors). The CQIN Bayesian network model is a risk exploration tool and predictive model for QI practitioners, fostering discussion of which features are important within a specific example, and allowing ‘what if’ analysis of potential interventions.

In the context of a network of adaptive agents delivering and improving healthcare services, ‘agents’ refers literally to everyone involved — clinicians, patients, families, management and support staff, suppliers of consumables and medical equipment, policy makers and funders. Boundaries between these agent groups are dynamic and subjective, but the complex adaptive system functions as a holistic system. CQIN elevates cooperation, shared understanding, and signal integration within a network of cooperating agents. This framing facilitates intra-team and inter-team conversations across the entire system, essential to the goal of building up the shared schema and cooperation necessary for complex change.

For QI practitioners, it is useful to understand the available complexity-informed improvement frameworks and which aspects of complexity they emphasise. There is no need to reinvent this work. However, it is equally useful to be able to respond to complexity with bespoke interventions built from consistent first principles, such as

CQIN. The identification of the CQIN strategies is more than an academic contribution. These strategies are practical because they can be deconstructed so easily. If our shared understanding is poor, we can devise interventions to identify the gaps and increase alignment. If our agent network structure is weak, practical steps can be taken to increase communication between communities within the agent network.

Practical QI interventions do not begin at the macro system level. CAS change is driven from agent activity and complexity-informed QI must still allow human agents to seek localised goal-driven improvement. Perhaps the most important contribution of this research is the demonstration of an emergent QI goal state from simple rules, grounded in the evolutionary foundations of CAS, within a relevant socio-technical context. The implications for QI methods and organisations are considerable. The conditions required for adaptation cannot be ignored and the time required for change is not related to externally imposed project milestones, however well planned. Path dependence (CAS history) will limit the available transformational options and change management must recognise the indivisibility of process and improvement. Expressed directly, a CAS interpretation of QI demands a very different approach to QI projects.

12.4. Limitations

This research has acknowledged limitations. An initial obstacle is primarily ontological. Readers may differ in agreement to the underlying (critical) realist ontology and the extent that natural science laws apply to human systems. The CAS interpretation is not being applied as a metaphor; it is proposed as an explanation. Advocates of human exceptionalism within the natural world may find the explanation unconvincing or insist that it remains a metaphor only. A second philosophical question is the risk of underdetermination. CQIN is one CAS theory of QI that has shown some support from the experimental and case data, but other theoretical explanations have not been comprehensively ruled out.

As a model of complexity, CQIN might be seen as an overly ‘simplified’ view of complexity. Unlike other approaches, it does not elaborate on structure, either in a detailed way such as with dynamic systems and causal loop modelling, or the more abstract but still formal structural approach of viable systems modelling. Healthcare professionals are described only as agents, not as individuals and not differentiated by their roles. The mitigation for these legitimate questions is based on the chosen strategy

for studying a socio-technical CAS. The research perspective is that the detailed structure of clinical microsystems and wider organisational contexts cannot be meaningfully reconstructed from a distance. Similarly, the psychology and capabilities of individual agents and their rich interactions are not easy to access or observe.

The response to these challenges was to concentrate on the *impact* of agent problem-solving interactions as the first level of enquiry, rather than the exhaustive details of these interactions. CQIN does not attempt to model the nuance of a specific healthcare QI endeavour. The objective was to create a proforma model that is consistent with theory where detail can be added as relevant within practical applications of the model.

The treatment of system feedback and emergent phenomena is restricted by the 'moment in time' view that CQIN provides. It is important to view any instance of the model as one static frame within a continuous picture. For example, if an improvement action leads to an unintended consequence, this consequence becomes a system feedback signal for the next learning cycle. Modelling at that level of detail requires a shift from the generic CQIN model to a specific case that can then be observed in full over time.

The unbalanced nature of the secondary data, due to reliance on published cases studies, has been addressed in the analysis but it is accepted that more work on *unsuccessful* QI is still required. A related issue arising from the use of secondary data is the definition and verification of QI 'success'. This study accepted the reports of published peer-reviewed articles at face value. Evidence of ongoing success and avoiding unintended consequences requires more longitudinal studies to be undertaken.

A final comment is also required on the accuracy of the relative influence applied to the CQIN constructs within the Bayesian Network model. The proportions observed from the secondary data apply for that data set only, and within the limits of the detection method (thematic coding by one researcher). Within these limits, saturation was achieved, and it is expected that replication would achieve similar results. Different methods of observation and detection may reveal different estimates of relative influence of the CQIN constructs. More generally, it is recognised that the empirical findings have been shaped by the underlying theory used to define and assess the CQIN model, and the practical findings are dependent on this theory. In the words of Werner Heisenberg:

"What we observe is not nature itself, but nature exposed to our method of questioning" (as quoted in Gleiser, 2018).

12.5. Recommendations for Future Research

CQIN is already being applied as a practical quality improvement tool by the author. In addition to this ongoing practical exploration, further research will help to refine and strengthen the initial work.

The Bayesian network model will be refined by adding new cases to the training data. To gather information from further cases, the survey as used in the applied case study (Appendix 6) can be easily adapted for data collection. Additionally, the relative influence of the constructs should be reviewed by a wider range of subject matter experts. A Delphi panel process will refine the model structure and the node probability expressions where judgement has been applied. Ideally, new QI improvement cases will include longitudinal studies and examples where QI has been judged to be unsuccessful.

Access to ongoing case studies will allow both modelling approaches used (ABM and BN) to move from generic to specific examples, and this in turn will allow better representation of the ongoing, dynamic changes within a healthcare QI complex adaptive system. Any of the CQIN constructs can also be studied in a greater level of granularity, with shared problem representation, integration of feedback signals, problem solving methods, and network centrality as obvious starting candidates. These are all active research areas in various knowledge domains; CQIN provides a template for bringing research on these elements together within an integrated QI theory.

The exploration of game-theoretic approaches to QI is also an obvious candidate for further research. There is much common ground in terms of cooperation, collective decision making, establishing shared problem representation, and updating agent schema with new information. Combining more detailed social network analysis with game theory focussed on collaborative QI problems is an intriguing possibility.

12.6. Concluding Thoughts

This study has asserted that interpreting QI as a CAS requires a mechanism for how CAS principles apply within socio-technical contexts and evidence for this explanation. An explanation developed from existing theory and evidence supporting the explanation have been provided. The central idea is perhaps hiding in plain sight within the literature. It is

not controversial to point out that humans, as a social species, cooperate to solve problems. Using cooperation as a measure of system order is fully compatible with biological, ecological and physics-based conceptions of natural CAS.

This research posits communication and cooperation within a network of adaptive agents as the essential element of a socio-technical CAS. Therefore, the network relationships and information sharing processes are critical to achieving QI outcomes. The creation of knowledge is a cultural process and knowledge is a cultural artefact. Although the exact details of how culture evolves are far from settled, the notion that we communicate to build and transfer knowledge in ways that go beyond genetic inheritance and direct imitation is not disputed. Unfortunately, our (necessary) socio-technical structures can constrain the essential pathways for this knowledge transference, and the presence of complexity factors increases these constraints. This research has highlighted the underlying evolutionary requirements for change and suggests that maintaining these requirements must take place alongside any other QI endeavours.

An evolutionary approach to quality improvement will require different ways of working and may be very difficult to achieve in practice. Yet the approach is worth considering seriously. Cooperating networks of adaptive agents can respond to the innumerable and unforeseen difficulties of complex systems, enabled or impaired by their socio-technical contexts. Any approach to transformational change in healthcare needs to include and nurture these simple, necessary prerequisites for change. Rather than inhibit change, the right organisational structures can also enable and amplify improvement ideas and capability. If this change in thinking can take place, then the possibility of quality improvement as an emergent CAS phenomenon becomes realistic. As per any CAS, this process does not begin at the higher levels of executive management or government policy. We start by sharing an idea and choosing to cooperate.

References

Adami, C. (2002). What is complexity? *BioEssays*, *24*(12), 1085–1094. https://doi.org/10.1002/bies.10192

Adami, C. (2012). The use of information theory in evolutionary biology. *Annals of the New York Academy of Sciences*, *1256*(1), 49-65.

Adler-Milstein, J., Nong, P., & Freidman, C.P. (2018). Preparing healthcare delivery organizations for managing computable knowledge. *Learning Health Sys*. *3*(2). https://doi.org/10.1002/lrh2.10070

Adolf, L., & Custer, H.H. (2002). Field notes from two trailbreakers: lessons learned using the CQIN trailblazer for self-assessment in S.E. Van Kollenberg (Ed.), *A collection of papers on self-study and institutional improvement,* 80-82.

Akhnif, E., Macq, J., Idrissi Fakhreddine, M. O., & Meessen, B. (2017). Scoping literature review on the Learning Organisation concept as applied to the health system. *Health research policy and systems*, *15*(1), 16. https://doi.org/10.1186/s12961-017-0176-x

Akmal, A., Podgorodnichenko, N., Foote, J., Greatbanks, R., Stokes, T., & Gauld, R. (2021). Why is quality improvement so challenging? A viable systems model perspective to understand the frustrations of healthcare quality improvement managers. *Health policy*, *125*(5), 658-664. https://doi.org/10.1016/j.healthpol.2021.03.015

Alhaider, A. A., Lau, N., Davenport, P. B., & Morris, M. K. (2020). Distributed situation awareness: a health-system approach to assessing and designing patient flow management. *Ergonomics*, *63*(6), 682–709. https://doi.org/10.1080/00140139.2020.1755061

Altman, L., Woolfenden, S., & Breen, C. (2017b). Kids GPS integrated care - leaving the Ivory Tower. *International Journal of Integrated Care*, *17*(3), 29. https://doi.org/10.5334/ijic.3141

Anderson, J. R. (1993). Problem solving and learning. *American psychologist*, *48*(1), 35.

Anderson, J. R. (2005). *Cognitive psychology and its implications.* (6th ed.). Worth Publishers.

Appelbaum, S. H. (1997). Socio-technical systems theory: an intervention strategy for organizational development. *Management Decision*, *35*(6), 452–463. https://doi.org/10.1108/00251749710173823

Aragon, T. J. (2017). PDSA Problem Solving: with a gentle introduction to double-loop learning, program theory, and causal graphs. https://escholarship.org/content/qt8wp451vd/qt8wp451vd.pdf

Argyris, C. (1976). Single-Loop and Double-Loop Models in Research on Decision Making. *Administrative Science Quarterly*, *21*(3), 363–375. https://doi.org/10.2307/2391848

Arksey, H., & O'Malley, L. (2005). Scoping Studies: Towards a Methodological Framework. *International Journal of Social Research Methodology*, 8(1), 19–32. https://doi.org/10.1080/1364557032000119616

Ashby, W. Ross. (1956). *An introduction to cybernetics*. Chapman and Hall.

Augustsson, H., Churruca, K., & Braithwaite, J. (2020). Change and improvement 50 years in the making: a scoping review of the use of soft systems methodology in healthcare. *BMC health services research* *20*(1), 1-13. https://doi.org/10.1186/s12913-020-05929-5

Backhouse, A., & Ogunlayi, F. (2020). Quality improvement into practice. The BMJ, m865. https://doi.org/10.1136/bmj.m865

Badham, J., Chattoe-Brown, E., Gilbert, N., Chalabi, Z., Kee, F., & Hunter, R. F. (2018). Developing agent-based models of complex health behaviour. *Health & place*, *54*, 170-177. https://doi.org/10.1016/j.healthplace.2018.08.022

Bain, R. (2016). Are our brains Bayesian? *Significance*, *13*(4), 14-19. https://doi.org/10.1111/j.1740-9713.2016.00935.x

Bak, P., & Chen, K. (1991) Self-organised criticality, *Scientific American*, *264* (1), 46-53.

Barabási, A-L. (2016). *Network Science*. Cambridge University Press.

Barach, P & Kleinman, L. (2018) Measuring and improving comprehensive pediatric cardiac care: Learning from continuous quality improvement methods and tools, *Progress in Pediatric Cardiology*, *48*, 82-92. https://doi.org/10.1016/j.ppedcard.2018.02.010

Barson, S., Doolan-Noble, F., Gray, J. A., & Gauld, R. (2017). Healthcare leaders' views on successful quality improvement initiatives and context. *Journal of Health Organisation and Management*, *31*(1), 54–63. https://doi.org/10.1108/jhom-10-2016-0191

Bar-Yam, Y. (2004). Multiscale variety in complex systems. *Complexity 9* (4), 37-45. https://doi.org/10.1002/cplx.20014

Bashan, A., Bartsch, R. P., Kantelhardt, J. W., Havlin, S., & Ivanov, P. C. (2012). Network physiology reveals relations between network topology and physiological function. *Nature Communications*, *3*(1). https://doi.org/10.1038/ncomms1705

Batalden, P. B., & Davidoff, F. (2007). What is "quality improvement" and how can it transform healthcare?. *Quality & safety in health care*, *16*(1), 2–3. https://doi.org/10.1136/qshc.2006.022046

Bauer, M. S., Damschroder, L., Hagedorn, H., Smith, J., & Kilbourne, A. M. (2015). An introduction to implementation science for the non-specialist. *BMC psychology*, *3*(1). https://doi.org/10.1186/s40359-015-0089-9

Beckford, J. (2002). *Quality (2nd Ed.).* Routledge.

Beer, S. (1984). The Viable System Model: its provenance, development, methodology and pathology. *Journal of the Operational Research Society*, *35*(1), 7–25. https://doi.org/10.1057/jors.1984.2

Berg-Schlosser, D., De Meur, G., Ragin, C. C., & Rihoux, B. (2009). Qualitative Comparative Analysis (QCA) as an approach. In *SAGE Publications, Inc. eBooks* (pp. 1–18). https://doi.org/10.4135/9781452226569.n1

Best, A., Greenhalgh, T., Lewis, S., Saul, J. E., Carroll, S., & Bitz, J. (2012). Large-system transformation in health care: a realist review. *The Milbank Quarterly*, *90*(3), 421-456. https://doi.org/10.1111/j.1468-0009.2012.00670.x

Birkenhäger-Gillesse E.G., Kollen, B.J., Achterberg, W.P., Boersma, F., Jongman, L., & Zuidema, S.U. (2018). Effects of Psychosocial Interventions for Behavioral and Psychological Symptoms in Dementia on the Prescription of Psychotropic Drugs: A Systematic Review and Meta-Analyses. *Journal of the American Medical Directors Association 19* (3), https://doi.org/10.1016/j.jamda.2017.12.100

Bishop, S., & Waring, J. (2019). From boundary object to boundary subject; the role of the patient in coordination across complex systems of care during hospital discharge. *Social Science & Medicine*, *235*, 112370. https://doi.org/10.1016/j.socscimed.2019.112370

Blackman, D., Buick, F., O'Donnell, M., & Ilahee, N. (2022). Changing the conversation to create organizational change. *Journal of Change Management*, *22*(3), 252–272. https://doi.org/10.1080/14697017.2022.2040570

Blaikie, Norman (2000). Chapter 3, "Research questions and objectives" in *Designing social research: the logic of anticipation*. Polity Press.

Booth, B., Zwar, N., & Harris, M. (2013). Healthcare improvement as planned system change or complex responsive processes? a longitudinal case study in general practice. *BMC Family Practice*, *14*(1). https://doi.org/10.1186/1471-2296-14-51

Borrill, P., & Tesfatsion, L. (2011). Agent-based modelling: the right mathematics for the social sciences? In J.B. Davis & D.W. Hands (Eds.), *Elgar Companion to Recent Economic Methodology*. Edward Elgar Publishers.

Braithwaite, J. (2018). Changing how we think about healthcare improvement. *The BMJ*, k2014. https://doi.org/10.1136/bmj.k2014

Braithwaite, J., Churruca, K., Long, J.C., Ellis, L.A., & Herkes, J. (2018). When complexity science meets implementation science: a theoretical and empirical analysis of systems change. *BMC Medicine 16* (1). https://doi.org/10.1186/s12916-018-1057-z

Braithwaite, J., Marks, D., & Taylor, N. (2014). Harnessing implementation science to improve care quality and patient safety: a systematic review of targeted literature. *International Journal for Quality in Health Care*, *26*(3), 321-329. https://doi.org/10.1093/intqhc/mzu047

Brandão de Souza, L. & Pidd, M. (2011). Exploring the barriers to lean health care implementation, *Public Money & Management*, *31*(1), 59-66. https://doi.org/10.1080/09540962.2011.545548

Britto, M. T., Fuller, S., Kaplan, H. C., Kotagal, U. R., Lannon, C., Margolis, P. A., Muething, S. E., Schoettker, P. J., & Seid, M. (2018b). Using a network organisational architecture to support the development of Learning Healthcare Systems. *BMJ Quality & Safety*, *27*(11), 937–946. https://doi.org/10.1136/bmjqs-2017-007219

Bryman, A. M. (2012). *Social Research Methods (4th Ed.)*. Oxford University Press.

Brys, N., Keating, J. A., Knobloch, M. J., & Safdar, N. (2019). Engaging patients in health care epidemiology research: A case example. *American Journal of Infection Control*, *47*(2), 139–143. https://doi.org/10.1016/j.ajic.2018.08.013

Buckley, W. (2017). Society as a complex adaptive system. In *Systems research for behavioral science* (pp. 490-513). Routledge.

Burns, R. B. (2000). Ethics of Research, chapter 2 in *Introduction to Research Methods*, Longman Publishing.

Callahan, C. M., Boustani, M., Weiner, M., Beck, R. A., Livin, L. R., Kellams, J. J., Willis, D. R., & Hendrie, H. C. (2010b). Implementing dementia care models in primary care settings: The Aging Brain Care Medical Home. *Aging & Mental Health*, *15*(1), 5–12. https://doi.org/10.1080/13607861003801052

Campbell, D. T. (1971). Reforms as Experiments. *Urban Affairs Quarterly*, *7*(2), 133–171. https://doi.org/10.1177/107808747100700202

Carmichael, T., & Hadžikadić, M. (2019). The fundamentals of complex adaptive systems. In *Understanding complex systems* (pp. 1–16). https://doi.org/10.1007/978-3-030-20309-2_1

Carney, T. J., Morgan, G. P., Jones, J., McDaniel, A. M., Weaver, M. T., Weiner, B. J., & Haggstrom, D. A. (2015). Hypothesis generation using network structures on community health center cancer-screening performance. *Journal of Biomedical Informatics*, *57*, 288–307. https://doi.org/10.1016/j.jbi.2015.08.005

Carter, J., Dwyer, N., Roselund, J., Cote, K., Chase, P., Martorana, K., Gustafson, K. E., Bergling, E., Kamalia, R., & DeGrazia, M. (2017). Systematic changes to help parents of medically complex infants manage medical expenses. *Advances in Neonatal Care*, *17*(6), 461–469. https://doi.org/10.1097/anc.0000000000000447

Castellani, B., & Gerrits, L. (2021). Map of the complexity sciences. Art and Science Factory, LLC. https://www.art-sciencefactory.com/complexity-map_feb09.html [Retrieved 19/07/2022]

Chandler, J., Rycroft-Malone, J., Hawkes, C., & Noyes, J. (2015). Application of simplified Complexity Theory concepts for healthcare social systems to explain the implementation of evidence into practice. *Journal of Advanced Nursing*, *72*(2), 461–480. https://doi.org/10.1111/jan.12815

Checkland, P. (2000). Soft systems methodology: a thirty year retrospective. Systems Research and Behavioral Science, 17(S1), S11–S58. https://doi.org/10.1002/1099-1743(200011)17:1

Clancy, T. R. (2009). Putting it altogether. *JONA: The Journal of Nursing Administration*, *39*(9), 364–367. https://doi.org/10.1097/nna.0b013e3181b3b6a8

Constantinou, A. C., & Fenton, N. (2018). Things to know about Bayesian Networks: Decisions under Uncertainty, Part 2. *Significance*, *15*(2), 19–23. https://doi.org/10.1111/j.1740-9713.2018.01126.x

Covvey, H. D. (2018). Healthcare as a complex adaptive system. In *EHealth research, theory and development* (pp. 69-90). Routledge.

Creanza, N., Kolodny, O., & Feldman, M. W. (2017). Cultural evolutionary theory: How culture evolves and why it matters. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(30), 7782–7789. https://doi.org/10.1073/pnas.1620732114

Cresswell, K. M., Worth, A., & Sheikh, A. (2010). Actor-Network Theory and its role in understanding the implementation of information technology developments in healthcare. *BMC medical informatics and decision making*, *10*(1), 1-11. https://doi.org/10.1186/1472-6947-10-67

Cribb, A., Entwistle, V., & Mitchell, P. (2022). Talking it better: conversations and normative complexity in healthcare improvement. *Medical humanities, 48*(1), 85-93. https://doi.org/10.1136/medhum-2020-012129

Cronin, M. A., & Weingart, L. R. (2007). Representational gaps, information processing, and conflict in functionally diverse teams. *Academy of Management Review, 32*(3), 761–773. https://doi.org/10.5465/amr.2007.25275511

Cronqvist, L. (2019). *Tosmana* [Version 1.61]. University of Trier. Internet: https://www.tosmana.net

Cunningham, F. C., Ranmuthugala, G., Plumb, J., Georgiou, A., Westbrook, J. I., & Braithwaite, J. (2012). Health professional networks as a vector for improving healthcare quality and safety: a systematic review. *BMJ quality & safety, 21*(3), 239-249. 239–249. https://doi.org/10.1136/bmjqs-2011-000187

Czulada, L., Leino, P., & Willis, T. S. (2016). Partnering with a family advisor to improve communication in a pediatric intensive care unit. *American Journal of Medical Quality, 31*(6), 520–525. https://doi.org/10.1177/1062860615600858

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation science, 4*(1), 1-15.

Dandoy, C. E., Hausfeld, J., Flesch, L., Hawkins, D., Demmel, K., Best, D., Osterkamp, E., Bracke, T., Nagarajan, R., Jodele, S., Holt, J., Giaccone, M. J., Davies, S. M., Kotagal, U. R., & Simmons, J. M. (2015). Rapid cycle development of a multifactorial intervention achieved sustained reductions in central line-associated bloodstream infections in haematology oncology units at a children's hospital: a time series analysis. *BMJ Quality & Safety, 25*(8), 633–643. https://doi.org/10.1136/bmjqs-2015-004450

D'Andreamatteo, A., Ianni, L., Lega, F., & Sargiacomo, M. (2015). Lean in healthcare: A comprehensive review. *Health Policy, 119*(9), 1197–1209. https://doi.org/10.1016/j.healthpol.2015.02.002

Davidoff, F., Dixon-Woods, M., Leviton, L. C., & Michie, S. (2015). Demystifying theory and its use in improvement. *BMJ Quality & Safety*, *24*(3), 228–238. https://doi.org/10.1136/bmjqs-2014-003627

Davis, J.P., Eisenhardt, K.M., and Bingham, C.B. (2007). Developing theory through simulation methods. *The Academy of Management Review*, *32*, (2), 480-499.

Dawkins, R. (1976) *The selfish gene*. Oxford University Press.

Deming, W. E. (1986). *Out of the Crisis.* MIT Press.

Deming, W. E. (2000). *The new economics: For industry, government, education.* MIT Press.

Dennett, D. C. (2017). *From bacteria to Bach and back: The evolution of minds*. WW Norton & Company.

Deutsch, D. (2011). *The beginning of infinity: explanations that transform the world*. Penguin UK.

Dixon-Woods, M., & Martin, G. (2016). Does quality improvement improve quality? *Future Hospital Journal, 3*(3), 191–194. https://doi.org/10.7861/futurehosp.3-3-191

Dooley, K. (1997). A complex adaptive systems model of organisational change. *Nonlinear Dynamics, Psychology and Life Sciences*, *1* (1), 69-97.

Doolittle, P. (2014). Complex constructivism: A theoretical model of complexity and cognition. *International Journal of Teaching and Learning in Higher Education*, *26* (3). 485-498.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from Cases: Opportunities and challenges. *Academy of Management Journal*, *50*(1), 25–32. https://doi.org/10.5465/amj.2007.24160888

Elertson, K. M. (2017). Improving completion of the initial comprehensive health assessment for children entering foster care. *Children and Youth Services Review*, *81*, 124–128. https://doi.org/10.1016/j.childyouth.2017.08.006

Erdek, M., & Pronovost, P. J. (2004). Improving assessment and treatment of pain in the critically ill. *International Journal for Quality in Health Care*, *16*(1), 59–64. https://doi.org/10.1093/intqhc/mzh010

Essén, A., & Lindblad, S. (2013). Innovation as emergence in healthcare: Unpacking change from within. *Social Science & Medicine*, *93*, 203–211. https://doi.org/10.1016/j.socscimed.2012.08.035

Fenton, N., & Neil, M. (2019). *Risk assessment and decision analysis with Bayesian networks* (2nd. ed.). CRC Press.

Fenton, N. E., Noguchi, T., & Neil, M. (2019). An extension to the noisy-OR function to resolve the 'explaining away' deficiency for practical Bayesian network problems. *IEEE transactions on knowledge and data engineering*, *31*(12), 2441-2445.

Ferreira, D. M. C. & T. A. Saurin, T.A. (2019). A complexity theory perspective of kaizen: a study in healthcare. *Production Planning & Control 30*(16): 1337-1353.

Filho, E. and Tenenbaum, G. (2020). Team Mental Models. In G. Tenenbaum & R.C. Eklund (Eds) *Handbook of Sport Psychology*. Wiley https://doi.org/10.1002/9781119568124.ch29

Fitzner, K. (2007). Reliability and validity a quick review. *The Diabetes Educator*, *33*(5), 775-780.

Fjeldstad, Ø. D., Snow, C. C., Miles, R. E., & Lettl, C. (2012). The architecture of collaboration. *Strategic Management Journal*, *33*(6), 734–750. https://doi.org/10.1002/smj.1968

Flood, R. L. (1998). "Fifth Discipline": Review and Discussion. *Systemic practice and action research*, *11*, 259-273.

French, S. (2015). Cynefin: uncertainty, small worlds and scenarios. *Journal of the Operational Research Society*, *66*(10), 1635–1645. https://doi.org/10.1057/jors.2015.21

Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nature reviews neuroscience*, *11*(2), 127-138.

Fryer, T. (2020). *A short guide to ontology and epistemology: why everyone should be a critcal realist.* Tom Fryer, 2020 Fryer.com/ontology-guide (retrieved 14/07/2022) https://tfryercom.files.wordpress.com/2020/10/cr_shortguide_201029.pdf

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models.* John Wiley & Sons.

Gershenson, C. (2013). The implications of interactions for science and philosophy. *Foundations of Science, 18*, 781-790.

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice versa). In *Subjective probability* (pp. 129-161). Wiley.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, *103*(4), 650.

Gleiser, M. (2018). How much can we know? *Nature*, *557*(7706), S20-S20. https://doi.org/10.1038/d41586-018-05100-5

Goderis, G., Borgermans, L., Mathieu, C., Van Den Broeke, C., Hannes, K., Heyrman, J., & Grol, R. (2009). Barriers and facilitators to evidence based care of type 2 diabetes patients: experiences of general practitioners participating to a quality improvement program. *Implementation Science*, *4*(1). https://doi.org/10.1186/1748-5908-4-41

Goldspink, C., & Kay, R. (2003). Organizations as self-organizing and sustaining systems: a complex and autopoietic systems perspective. *International Journal of General Systems, 32*(5), 459-474.

Goodyear-Smith, F., & Ashton, T. (2019). New Zealand health system: universalism struggles with persisting inequities. *The Lancet, 394*(10196), 432-442.

Gorski, P. S. (2013). "What is Critical Realism? And Why Should You Care?" *Contemporary Sociology*, *42*(5), 658–670. https://doi.org/10.1177/0094306113499533

Greenhalgh, T., Papoutsi, C. (2018). Studying complexity in health services research: desperately seeking an overdue paradigm shift. *BMC Med 16*, 95. https://doi.org/10.1186/s12916-018-1089-4

Griffin, A., McKeown, A., Viney, R., Rich, A., Welland, T., Gafson, I., & Woolf, K. (2017). Revalidation and quality assurance: the application of the MUSIQ framework in independent verification visits to healthcare organisations. *BMJ Open, 7*(2), e014121. https://doi.org/10.1136/bmjopen-2016-014121

Guba, E.G. (Ed.). (1990). *The paradigm dialog.* Sage.

Hanckel, B., Petticrew, M., Thomas, J., & Green, J. (2021). The use of Qualitative Comparative Analysis (QCA) to address causality in complex systems: a systematic review of research on public health interventions. *BMC Public Health, 21*(1). https://doi.org/10.1186/s12889-021-10926-2

Haraldseid-Driftland, C., Billett, S., Guise, V., Schibevaag, L., Alsvik, J. G., Fagerdal, B., ... & Wiig, S. (2022). The role of collaborative learning in resilience in healthcare—a thematic qualitative meta-synthesis of resilience narratives. *BMC Health Services Research, 22*(1), 1-12.

Hawe, P. (2015). Lessons from complex interventions to improve health. *Annual review of public health, 36*, 307-323.

Health and Disability Ethics Committees. (2018). HDEC – Scope of Review Form. https://ethics.health.govt.nz.

Heifetz, R.A., Heifetz, R., Grashow, A., and Linsky, M. (2009). *The practice of adaptive leadership: Tools and tactics for changing your organization and the world.* Harvard Business Press.

Hignett, S., Wolf, L., Taylor, E., & Griffiths, P. (2015). Firefighting to innovation. *Human Factors, 57*(7), 1195–1207. https://doi.org/10.1177/0018720815593642

Hilts, L., Howard, M., Price, D., Risdon, C., Agarwal, G., & Childs, A. M. (2012). Helping primary care teams emerge through a quality improvement program. *Family Practice, 30*(2), 204–211. https://doi.org/10.1093/fampra/cms056

Holland, J. H. (1992). Complex adaptive systems. *Daedalus, 121*(1), 17-30.

Holland, J. H. (2006). Studying complex adaptive systems. *Journal of Systems Science and Complexity, 19*(1), 1–8. https://doi.org/10.1007/s11424-006-0001-z

Holland, J. H. (2012). *Signals and boundaries: Building blocks for complex adaptive systems*. MIT Press.

Hollnagel, E., Wears, R. L., & Braithwaite, J. (2015). *From Safety-I to Safety-II: a white paper*. The resilient health care net: published simultaneously by the University of Southern Denmark, University of Florida, USA, and Macquarie University, Australia.

Horn, M., & Gorman, D. (2021). The political economy of healthcare reform: why New Zealand has experienced 82 years of ineffectual reforms and what can be done about it. *The New Zealand Medical Journal (Online), 134*(1533), 104-7.

Hoyle, E., Patino, F., & Yoxall, C. W. (2019). Quality improvement programme to improve compliance with initial respiratory support guideline at preterm birth. *Acta Paediatrica*, *109*(5), 943–947. https://doi.org/10.1111/apa.15076

Hüllermeier, E., & Waegeman, W. (2021). Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Machine Learning, 110*(3), 457–506. https://doi.org/10.1007/s10994-021-05946-3

Imai, M. (2012). *Gemba Kaizen: A commonsense approach to a continuous improvement strategy*. McGraw Hill.

Institute for Healthcare Improvement (2023) https://www.ihi.org/resources/Pages/HowtoImprove/default.aspx

Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st century*. National Academy Press.

Jackson, M.C. (2001). Critical systems thinking and practice, *European Journal of Operational Research, 128*(2), 233-244.

Jackson, M. (2006). Creative holism: a critical systems approach to complex problem situations. *Systems Research and Behavioral Science, 23*(5), 647–657. https://doi.org/10.1002/sres.799

Jackson, M. (2020). How We Understand “Complexity” Makes a Difference: Lessons from Critical Systems Thinking and the Covid-19 Pandemic in the UK. *Systems, 8*(4), 52. https://doi.org/10.3390/systems8040052

Joffe, H. (2011). Thematic analysis. *Qualitative research methods in mental health and psychotherapy: A guide for students and practitioners*, 209-223.

Junghans, T. (2017). “Don’t mind the gap!” Reflections on improvement science as a paradigm. *Health Care Analysis*, *26*(2), 124–139. https://doi.org/10.1007/s10728-017-0353-7

Kaplan, H. C., Provost, L., Froehle, C. M., & Margolis, P. A. (2011). The Model for Understanding Success in Quality (MUSIQ): building a theory of context in healthcare quality improvement. *BMJ Quality & Safety, 21*(1), 13–20. https://doi.org/10.1136/bmjqs-2011-000010

Kauffman, S. (2000). *Investigations*. Oxford University Press.

Khan, S., Vandermorris, A., Shepherd, J., Begun, J., Lanham, H. J., Uhl-Bien, M, & Berta, W. (2018). Embracing uncertainty, managing complexity: applying complexity thinking principles to transformation efforts in healthcare systems. *BMC Health Services Research 18*(1). https://doi.org/10.1186s12913-01802994-0

Kitson, A., Brook, A., Harvey, G., Jordan, Z., Marshall, R., O’Shea, R., & Wilson, D. (2018). Using complexity and network concepts to inform healthcare knowledge translation. *International journal of health policy and management, 7*(3), 231.

Korb, K. B., Geard, N., & Dorin, A. (2013). A Bayesian Approach to the Validation of Agent-Based Models. In A. Tolk (Ed.), *Ontology, Epistemology, and Teleology for Modeling and Simulation: Philosophical Foundations for Intelligent M&S Applications* (pp. 255-269). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-31140-6_14

Kovach, J. V., & Fredendall, L. D. (2013). The influence of continuous improvement practices on learning: an empirical study. *Quality Management Journal*, *20*(4), 6–20. https://doi.org/10.1080/10686967.2013.11918361

Krakauer, D. C. (2023). Symmetry–simplicity, broken symmetry–complexity. *Interface Focus*, *13*(3). https://doi.org/10.1098/rsfs.2022.0075

Kringos, D., Sunol, R., Wagner, C., Mannion, R., Michel, P.,.Klazinga, N., Groene, O.,& on behalf of the DUQuE Consortium (2015).The influence of context on the effectiveness of hospital quality improvement strategies: a review of systematic reviews. *BMC Health Services Research 15*, 1-13. https://doi.org/10.1186/s12913-015-0906-0

Kurtz, C. F., & Snowden, D. J. (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM Systems Journal*, *42*(3), 462–483. https://doi.org/10.1147/sj.423.0462

Kyrimi, E., Neves, M. R., McLachlan, S., Neil, M., Marsh, W., & Fenton, N. (2020). Medical idioms for clinical Bayesian network development. *Journal of Biomedical Informatics*, *108*, 103495. https://doi.org/10.1016/j.jbi.2020.103495

Kyrimi, E., McLachlan, S., Dube, K., Neves, M. R., Fahmi, A., & Fenton, N. (2021). A comprehensive scoping review of Bayesian networks in healthcare: Past, present and future. *Artificial Intelligence in Medicine*, *117*, 102108. https://doi.org/10.1016/j.artmed.2021.102108

Lachman, P., Linkson, L., Evans, T., Clausen, H., & Hothi, D. K. (2015). Developing person-centred analysis of harm in a paediatric hospital: a quality improvement report. *BMJ Quality & Safety*, *24*(5), 337–344. https://doi.org/10.1136/bmjqs-2014-003795

Laubenbacher, R., Jarrah, A., Mortveit, H., & Ravi, S. (2007). A mathematical formalism for agent-based modelling. *arXiv:0801.0249v1 [cs.MA]*. [retrieved 23/11/2019].

Lemieux-Charles, L., Murray, M., Baker, G. R., Barnsley, J., Tasa, K., & Ibrahim, S. A. (2002). The effects of quality improvement practices on team effectiveness: a mediational model. *Journal of Organizational Behavior*, *23*(5), 533–553. https://doi.org/10.1002/job.154

Lemire, S., Christie, C. A., & Inkelas, M. (2017). The methods and tools of improvement science. In C. A. Christie, M. Inkelas & S. Lemire (Eds.), *Improvement Science in*

Evaluation: Methods and Uses: *New Directions for Evaluation, 153*. John Wiley & sons, 23– 33.

Leonard, A. (2009). The viable system model and its application to complex organizations. *Systemic practice and action research, 22*(4), 223-233.

Levula, A., Chung, K., Young, J., & White, K. (2013). Envisioning complexity in healthcare systems through social networks. *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 978-1-4503-2240-9. https://doi.org/10.1145/2492517.2500313

Lorden, A. L., Zhang, Y., Lin, S. H., & Côté, M. J. (2014). Measures of success: the role of human factors in lean implementation in healthcare. *Quality Management Journal*, *21*(3), 26-37.

Mahajan, A., Islam, S. D., Schwartz, M. J., & Cannesson, M. (2017). A hospital is not just a factory, but a complex adaptive system—implications for perioperative care. *Anesthesia & Analgesia,125*(1), 333-341.

Manukyan, N., Eppstein, M. J., & Horbar, J. D. (2013). Team Learning for Healthcare Quality Improvement. *IEEE Access*, *1*, 545–557. https://doi.org/10.1109/access.2013.2280086

Massey University. (2015). *Code of responsible research conduct* https://www.massey.ac.nz/massey/research/responsible-research-conduct/responsible-research-conduct_home.cfm [retrieved 06 February 2020].

Massey University. (2017). *Code of ethical conduct for research, teaching and evaluations involving human participants (revised code 2017).* https://www.massey.ac.nz/massey/research/research-ethics/research-ethics_home.cfm [retrieved 02 November 2019].

Maturana, H. R., & Varela, F. J. (1987). *The tree of knowledge: The biological roots of human understanding.* New Science Library/Shambhala Publications.

McLachlan, S., Potts, H. W. W., Dube, K., Buchanan, D., Lean, S., Gallagher, T., Johnson, O., Daley, B., Marsh, W., & Fenton, N. (2018). The Heimdall framework

for supporting characterisation of learning health systems. *BMJ Health & Care Informatics, 25*(2), 77–87. https://doi.org/10.14236/jhi.v25i2.996

McLachlan, S., Daley, B. J., Saidi, S., Kyrimi, E., Dube, K., Grosan, C., Neil, M., Rose, L., & Fenton, N. E. (2024). Approach and Method for Bayesian Network Modelling: A Case Study in Pregnancy Outcomes for England and Wales. *medRxiv*, 2024.2001.2006.24300925. https://doi.org/10.1101/2024.01.06.24300925

Meredith, J. (1993). Theory Building through Conceptual Methods. *International Journal of Operations & Production Management, Vol. (13)*5, 3-11. https://doi.org/10.1108/01443579310028120

Michel, P., Roberts, T., Porro, Z., Es-seddiqi, H., Saillour, F., et al. (2016) What are the barriers and facilitators to the implementation and/or success of quality improvement and risk management in hospitals: a systematic literature review. *J Epidemiol Public Health Rev 1*(4). https://doi.org/10.16966/2471-8211.123

Miller, John H., & Page, Scott E. (2007). *Complex Adaptive Systems: An introduction to computational models of social life*. Princeton University Press.

Mills, T., Lawton, R., & Sheard, L. (2019). Advancing complexity science in healthcare research: the logic of logic models. *BMC Medical Research Methodology, 19*(1). https://doi.org/10.1186/s12874-019-0701-4

Mitchell, M. (2009). *Complexity: A guided tour*. Oxford University Press.

Miton, H., & DeDeo, S. (2022). The cultural transmission of tacit knowledge. *Journal of the Royal Society Interface, 19*(195), 20220238.

Moen, R. D., & Norman, C. L. (2010). *Clearing up myths about the Deming cycle and seeing how it keeps evolving* https://apiweb.org/circling-back.pdf

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for Systematic Reviews and Meta-Analyses: the PRISMA statement. *PLoS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mohammed, S., Ferzandi, L., & Hamilton, K. (2010). Metaphor No More: A 15-Year Review of the Team Mental Model Construct. *Journal of Management, 36*(4), 876–910. https://doi.org/10.1177/0149206309356804

Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(20), 7176–7184. https://doi.org/10.1073/pnas.1319946111

Morin, E. (2007). Restricted complexity, general complexity. In *Worldviews, science and us: Philosophy and complexity*. University of Liverpool, 5-29.

Munn, Z., Peters M., Stern, C., Tufanaru, C., McArthur, A., Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology 18*(1). https://doi.org/10.1186s12874-018-0611-x

Murray, M., Higuchi, K. S., Edwards, N., Greenough, M., & Hoogeveen, K. (2011). Practice change in Long-Term Care Facilities in Ontario: Views of managers and staff. *Geriatric Nursing*, *32*(6), 418–428. https://doi.org/10.1016/j.gerinurse.2011.08.001

Muttalib, F., Ballard, E., Langton, J., Malone, S., Fonseca, Y., Hansmann, A., Remy, K. E., Hovmand, P., & Doctor, A. (2021a). Application of systems dynamics and group model building to identify barriers and facilitators to acute care delivery in a resource limited setting. *BMC Health Services Research, 21*(1). https://doi.org/10.1186/s12913-020-06014-7

Myron, R., French, C. E., Sullivan, P. A., Sathyamoorthy, G., Barlow, J., & Pomeroy, L. (2017). Professionals learning together with patients: An exploratory study of a collaborative learning Fellowship programme for healthcare improvement. *Journal of Interprofessional Care*, *32*(3), 257–265. https://doi.org/10.1080/13561820.2017.1392935

Neil, M., Fenton, N., & Nielson, L. (2000). Building large-scale Bayesian networks. *Knowledge Engineering Review*, *15*(3), 257–284. https://doi.org/10.1017/s0269888900003039

NetLogo. (2023a). Publications. https://ccl.northwestern.edu/netlogo/references.shtml

NetLogo. (2023b). https://modelingcommons.org/browse/one_model/7245#

Ngo, J., Lau, D., Ploquin, J., Receveur, T., Stassen, K., & Del Castilho, C. (2022). Improving incident reporting among physicians at south health campus hospital. *BMJ Open Quality*, *11*(4), e001945. https://doi.org/10.1136/bmjoq-2022-001945

Nordin, A. (2017). Expressions of shared interpretations - Intangible outcomes of continuous quality improvement efforts in health- and elderly care (PhD dissertation, Jönköping University, School of Health and Welfare). Retrieved from https://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-37302

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science, 314*(5805), 1560-1563.

Ognyanova, K. (2016). Network Analysis and Visualization with R and igraph. In *NetSciX 2016 School of Code Workshop*. https://kateto.net/netscix2016

Ogrinc, G., Davies, L., Goodman, D., Batalden, P. B., Davidoff, F., & Stevens, D. P. (2015). SQUIRE 2.0 (Standards for Quality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. *American Journal of Critical Care*, *24*(6), 466–473. https://doi.org/10.4037/ajcc2015455

Oreskes, N., Shrader-Frechette, K., & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth sciences. *Science, 263*(5147), 641-646.

Ornstein, J. T., Hammond, R. A., Padek, M., Mazzucca, S., & Brownson, R. C. (2020). Rugged landscapes: complexity and implementation science. *Implementation Science, 15*, 1-9.

Pae Ora (Healthy Futures) Act 2022 No 30 (as at 13 January 2023), Public Act – New Zealand Legislation. https://www.legislation.govt.nz/act/public/2022/0030/latest/versions.aspx

Paley J. (2010). The appropriation of complexity thinking in health care. *J Health Serv Res Policy*, *15*(1), 59–61.

Paley, J., & Eva, G. (2011). Complexity theory as an approach to explanation in healthcare: A critical discussion. *International Journal of Nursing Studies, 48*(2), 269–279. https://doi.org/10.1016/j.ijnurstu.2010.09.012

Pearl, J. & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. Penguin Books.

Pearson, M., Hemsley, A., Blackwell, R. E., Pegg, L. E., & Custerson, L. (2017). Improving Hospital at Home for frail older people: insights from a quality improvement project to achieve change across regional health and social care sectors. *BMC Health Services Research, 17*(1). https://doi.org/10.1186/s12913-017-2334-9

Perla, R. J., Provost, L. P., & Parry, G. J. (2013). Seven propositions of the science of improvement: exploring foundations. *Quality Management in Healthcare, 22*(3), 170-186.

Pesec, M., Spigel, L., Granados, J. M. M., Bitton, A., Hirschhorn, L. R., Brizuela, J. a. J., Pignone, M., Del Rocío Sáenz, M., Schwarz, D., Del Carpio, O. V., Wilson, I. B., Méndez, E. Z., & Ratcliffe, H. L. (2021). Strengthening data collection and use for quality improvement in primary care: the case of Costa Rica. *Health Policy and Planning*, *36*(5), 740–753. https://doi.org/10.1093/heapol/czab043

Plsek, P. E., & Greenhalgh, T. (2001). Complexity science: The challenge of complexity in health care. *BMJ. 323*(7313), 625–628. https://doi.org/10.1136/bmj.323.7313.625

Poksińska, B., & Swartling, D. (2018). From successful to sustainable Lean production – the case of a Lean Prize Award Winner. *Total Quality Management & Business Excellence*, *29*(9–10), 996–1011. https://doi.org/10.1080/14783363.2018.1486539

Praetorius, T., Chaudhuri, A., Saddikuti, V., Hasle, P., & Singh, A. (2018). Achieving better integration in trauma care delivery in India. *Journal of Health Management*, *20*(3), 234–254. https://doi.org/10.1177/0972063418779896

Preiser, R., Biggs, R., De Vos, A., & Folke, C. (2018). Social-ecological systems as complex adaptive systems: organizing principles for advancing research methods and approaches. *Ecology and Society, 23*(4). https://doi.org/10.5751/es-10558-230446

Puga, J., Krzywinski, M. & Altman, N. (2015). Bayesian networks. *Nature Methods, 12* (9), 799–800. https://doi.org/10.1038/nmeth.3550

Quinn, D., Jacobson, J. O., Wong, S. L., Wollins, D. S., Gilmore, T., & Cunningham, G. (2021). Improving the Delivery of High-Quality Cancer Care in Medically Underserved Communities: a formative Evaluation method. *Quality Management in Health Care*, *30*(4), 251–258. https://doi.org/10.1097/qmh.0000000000000313

Radcliffe, E., Kordowicz, M., Mak, C., Shefer, G., Armstrong, D., White, P., & Ashworth, M. (2020). Lean implementation within healthcare: imaging as fertile ground. *Journal of Health Organisation and Management*, *34*(8), 869–884. https://doi.org/10.1108/jhom-02-2020-0050

Ragin, C. C. (1999). Using qualitative comparative analysis to study causal complexity. *PubMed*, *34*(5 Pt 2), 1225–1239. https://pubmed.ncbi.nlm.nih.gov/10591281

Railsback, S. F., Lytinen, S. L., & Jackson, S. K. (2006). Agent-based simulation platforms: Review and development recommendations. *Simulation*, *82*(9), 609-623.

Rancati, A., & Snowden, D. (2021). *Managing complexity (and chaos) in times of crisis : a field guide for decision makers inspired by the Cynefin framework*, European Commission, Joint Research Centre Publications Office, 2021, https://data.europa.eu/doi/10.2760/353

Reed, J.E., & Card, A.J. (2016). The problem with Plan-Do-Study-Act cycles. *BMJ quality & safety*, *25*(3), 147–152. https://doi.org/10.1136/bmjqs-2015-005076

Reed, J. E., Howe, C., Doyle, C., & Bell, D. (2018). Simple rules for evidence translation in complex systems: a qualitative study. *BMC medicine*, *16*, 1-20.

Reed, J.E., Kaplan, H.C., & Ismail, S.A. (2018). A new typology for understanding context: qualitative exploration of the model for understanding success in quality (MUSIQ). *BMC Health Services Research*, *18*, 1-14.

Rescher, N. (1998). *Complexity: A philosophical overview*. Transaction Publishers.

Rihoux, B., & De Meur, G. (2009). Crisp-set qualitative comparative analysis (csQCA). In *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*, 33-68.

Riordan, F., McGrath, N., Dinneen, S. F., Kearney, P. M., & McHugh, S. (2019). 'Sink or Swim': A qualitative study to understand how and why nurses adapt to support

the implementation of integrated diabetes care. *International Journal of Integrated Care, 19*(2), 2. https://doi.org/10.5334/ijic.4215

Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy sciences, 4*(2), 155-169.

Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation, 14*(1), 29-48.

Rohrbasser, A., Harris, J., Mickan, S., Tal, K., & Wong, G. (2018). Quality circles for quality improvement in primary health care: Their origins, spread, effectiveness and lacunae– A scoping review. *PloS One, 13*(12), e0202616. https://doi.org/10.1371/journal.pone.0202616

Rubinson, C. (2019). Presenting qualitative comparative analysis: Notation, tabular layout, and visualization. *Methodological Innovations, 12*(2), 2059799119862110.

Sahin, O., Salim, H., Suprun, E., Richards, R., MacAskill, S., Heilgeist, S., … Beal, C. D. (2020). Developing a Preliminary Causal Loop Diagram for Understanding the Wicked Complexity of the COVID-19 Pandemic. *Systems, 8*(2), 20. https://doi.org/10.3390/systems8020020

Saurin, T. A., Rooke, J., & Koskela, L. (2013). A complex systems theory perspective of lean production. *International Journal of Production Research, 51*(19), 5824–5838. https://doi.org/10.1080/00207543.2013.796420

Scatà, M., Di Stefano, A., La Corte, A., Liò, P., Catania, E., Guardo, E., & Pagano, S. (2016). Combining evolutionary game theory and network theory to analyze human cooperation patterns. *Chaos, Solitons & Fractals, 91*, 17-24.

Schill, C., Anderies, J. M., Lindahl, T., Folke, C., Polasky, S., Cárdenas, J. C., ... & Schlüter, M. (2019). A more dynamic understanding of human behaviour for the Anthropocene. *Nature Sustainability, 2*(12), 1075-1082.

Schouten, L., Hulscher, M., Van Everdingen, J., Huijsman, R., & Grol, R. (2008). Evidence for the impact of quality improvement collaboratives: systematic review. *BMJ, 336*(7659), 1491–1494. https://doi.org/10.1136/bmj.39570.749884.be

Seid, M., Hartley, D., & Margolis, P. A. (2021). A science of collaborative learning health systems. *Learning Health Systems, 5*(3). https://doi.org/10.1002/lrh2.10278

Seid, M., Bridgeland, D. M., Bridgeland, A., & Hartley, D. (2021b). A collaborative learning health system agent-based model: Computational and face validity. *Learning Health Systems, 5*(3). https://doi.org/10.1002/lrh2.10261

Selznick, P. (1948). Foundations of the theory of organization. *American sociological review, 13*(1), 25-35.

Senge, P.M. (2006). *The fifth discipline: The art and practice of the learning organisation*. Random house.

Sharma, K., Jones, P., Cumming, J., & Middleton, L. (2024). Key elements and contextual factors that influence successful implementation of large-system transformation initiatives in the New Zealand health system: a realist evaluation. *BMC Health Services Research*, *24*(1). https://doi.org/10.1186/s12913-023-10497-5

Shelby, M., Malloch, K., Shellenberger, T. D., Byrum, S. D., Mackie, L., Dilli, S., & Martin, A. D. (2019). Addressing falls among older oncology patients through Complexity science. *Nursing Administration Quarterly*, *43*(3), 280–288. https://doi.org/10.1097/naq.0000000000000360

Simmons, M., Daniel, S., & Temple, M. (2019). How to manipulate friends and influence practice: Application of complexity science leads to quality improvement in laboratory sample submissions. *Journal of Infection Prevention, 20*(2), 91–98. https://doi.org/10.1177/1757177419831348

Simon, H. A. (2000). Bounded rationality in social science: Today and tomorrow. *Mind & Society, 1*, 25-39.

Small, M. L. (2009). 'How many cases do I need?': On science and the logic of case selection in field-based research. *Ethnography, 10*(1), 5–38. https://doi.org/10.1177/1466138108099586

Smerdon, D., Offerman, T., & Gneezy, U. (2016). *Everybody's doing it: On the Emergence and Persistence of Bad Social Norms* (No. 16-023/I). Tinbergen Institute Discussion Paper.

Smolla, M., Jansson, F., Lehmann, L., Houkes, W., Weissing, F. J., Hammerstein, P., ... & Enquist, M. (2021). Underappreciated features of cultural evolution. *Philosophical Transactions of the Royal Society B, 376*(1828), 20200259.

Soliman, M., & Saurin, T. A. (2017). Lean production in complex socio-technical systems: A systematic literature review. *Journal of Manufacturing Systems*, 45, 135-148.

Sun, Z., Lorscheid, I., Millington, J. D., Lauf, S., Magliocca, N. R., Groeneveld, J., ... & Buchmann, C. M. (2016). Simple or complicated agent-based models? A complicated issue. *Environmental Modelling & Software, 86*, 56-67.

Talbert-Slagle, K., Berg, D. N., & Bradley, E. H. (2013). Innovation spread: lessons from HIV. *International Journal for Quality in Health Care, 25*(4), 352–356. https://doi.org/10.1093/intqhc/mzt038

Tashakkori, A., & Teddlie, C. (1998). Chapter 3 in *Mixed methodology: combining qualitative and quantitative approaches*, Sage Publications.

Taylor, M., McNicholas, C., Nicolay, C., Darzi, A., Bell, D., & Reed, J. (2013). Systematic review of the application of the plan–do–study–act method to improve quality in healthcare. *BMJ Quality & Safety, 23*(4), 290–298. https://doi.org/10.1136/bmjqs-2013-001862

Thiele, J. C. (2014). R marries NetLogo: introduction to the RNetLogo package. *Journal of Statistical Software, 58*, 1-41.

Thompson, D. S., Fazio, X., Kustra, E., Patrick, L., & Stanley, D. (2016). Scoping review of complexity theory in health services research. *BMC Health Services Research, 16*(1). https://doi-org.ezproxy.massey.ac.nz/10.1186/s12913-016-1343-4

Tomasello, M. (2022). *The evolution of agency: Behavioural organization from lizards to humans*. MIT Press.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., ... & Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of internal medicine, 169*(7), 467-473.

Trist, E. L. (1981). *The evolution of socio-technical systems* (Vol. 2). Toronto: Ontario Quality of Working Life Centre.

Turner, J. R., & Baker, R. M. (2019). Complexity Theory: An Overview with Potential Applications for the Social Sciences. *Systems, 7*(1), 4. https://doi.org/10.3390/systems7010004

Uhl-Bien, M., Marion, R., & McKelvey, B. (2007). Complexity leadership theory: Shifting leadership from the industrial age to the knowledge era. *The leadership quarterly*, *18*(4), 298-318.

Uhl-Bien, M., & Marion, R. (2009). Complexity leadership in bureaucratic forms of organizing: A meso model. *The Leadership Quarterly*, *20*(4), 631-650.

Umpleby, S. A. (2009). Ross Ashby's general theory of adaptive systems. *International Journal of General Systems, 38*(2), 231-238. https://www.tandfonline.com/doi/abs/10.1080/03081070802601509.

Urbanowicz, R. J., & Moore, J. H. (2009). Learning Classifier Systems: A complete introduction, review, and roadmap. *Journal of Artificial Evolution and Applications*, *2009*, 1–25. https://doi.org/10.1155/2009/736398

Vaismoradi, M., & Snelgrove, S. (2019). Theme in Qualitative Content Analysis and Thematic Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 20*(3). https://doi.org/10.17169/fqs-20.3.3376

Van Leeuwen, E., Petrovic, M., van Driel, M., et al. (2018). Withdrawal versus continuation of long-term antipsychotic drug use for behavioural and psychological symptoms in older people with dementia. *Cochrane Database Syst Rev*, (3). https://doi.org/10.1002/14651858.CD007726.pub3

Vatnøy, T. K., Karlsen, T., & Dale, B. (2019). Exploring nursing competence to care for older patients in municipal in-patient acute care: A qualitative study. *Journal of Clinical Nursing, 28*(17–18), 3339–3352. https://doi.org/10.1111/jocn.14914

von Bertalanffy, L. (1972). The History and Status of General Systems Theory. *The Academy of Management Journal*, *15*(4), 407-426.

Voss, C., Tsikriktsis, N., & Frohlich, M. (2002) Case research in operations management. *International Journal of Operations & Production Management*, *22* (2), 195-219. https://doi.org/10.1108/01443570210414329

Wacker, J.G. (2008). A conceptual understanding of requirements for theory-building research: Guidelines for scientific theory building. *Journal of Supply Chain Management, 44*(3), 5-15. https://doi.org/10.1111/j.1745-493x.2008.00062.x

Walshe, K. (2007). Understanding what works—and why—in quality improvement: the need for theory-driven evaluation. *International Journal for Quality in Health Care*, *19*(2), 57—59.

Wang, S. C., Mannix, E. A., & Cronin, M. A. (2016). The influence of representational gaps on team processes and performance. Social Science Research Network. https://doi.org/10.2139/ssrn.2789851

Waring, J., & Bishop, S. (2010). Lean healthcare: Rhetoric, ritual and resistance. *Social Science & Medicine, 71*(7), 1332–1340. https://doi.org/10.1016/j.socscimed.2010.06.028

Weber, R. (2012). Evaluating and developing theories in the information systems discipline. *Journal of the Association for Information systems,13*(1), 2.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization science*, *16*(4), 409-421.

Wiesner, K., & Ladyman, J., (2019). Measuring complexity. *arXiv*:1909.13243.

Wilensky, U., & Rand, W. (2015). *An introduction to agent-based modelling: Modelling natural, social, and engineered complex systems with netlogo*. MIT Press.

Wilson, D. S., & Kirman, A. (Eds.). (2016). *Complexity and evolution: Toward a new synthesis for economics* (Vol. 19). MIT Press.

Wilson, W., McLachlan, S., Dube, K., Potter, K., & Jayamaha, N. (2023). Uncertainty, emergence and adaptation: A complex adaptive systems approach to quality improvement. *The Quality Management Journal*, *30*(3), 168–186. https://doi.org/10.1080/10686967.2023.2211287

Windrum, P., Fagiolo, G., & Moneta, A. (2007). Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation, 10*(2), 8.

Womack, JP. & Jones, DT. (2003). *Lean thinking: banish waste and create wealth in your corporation.* (2nd Ed.). Simon and Schuster.

Wynn, D.J., & Williams, C.K. (2012). Principles for conducting critical realist case study research in information systems. *MIS Quarterly, 36*(3), 787.

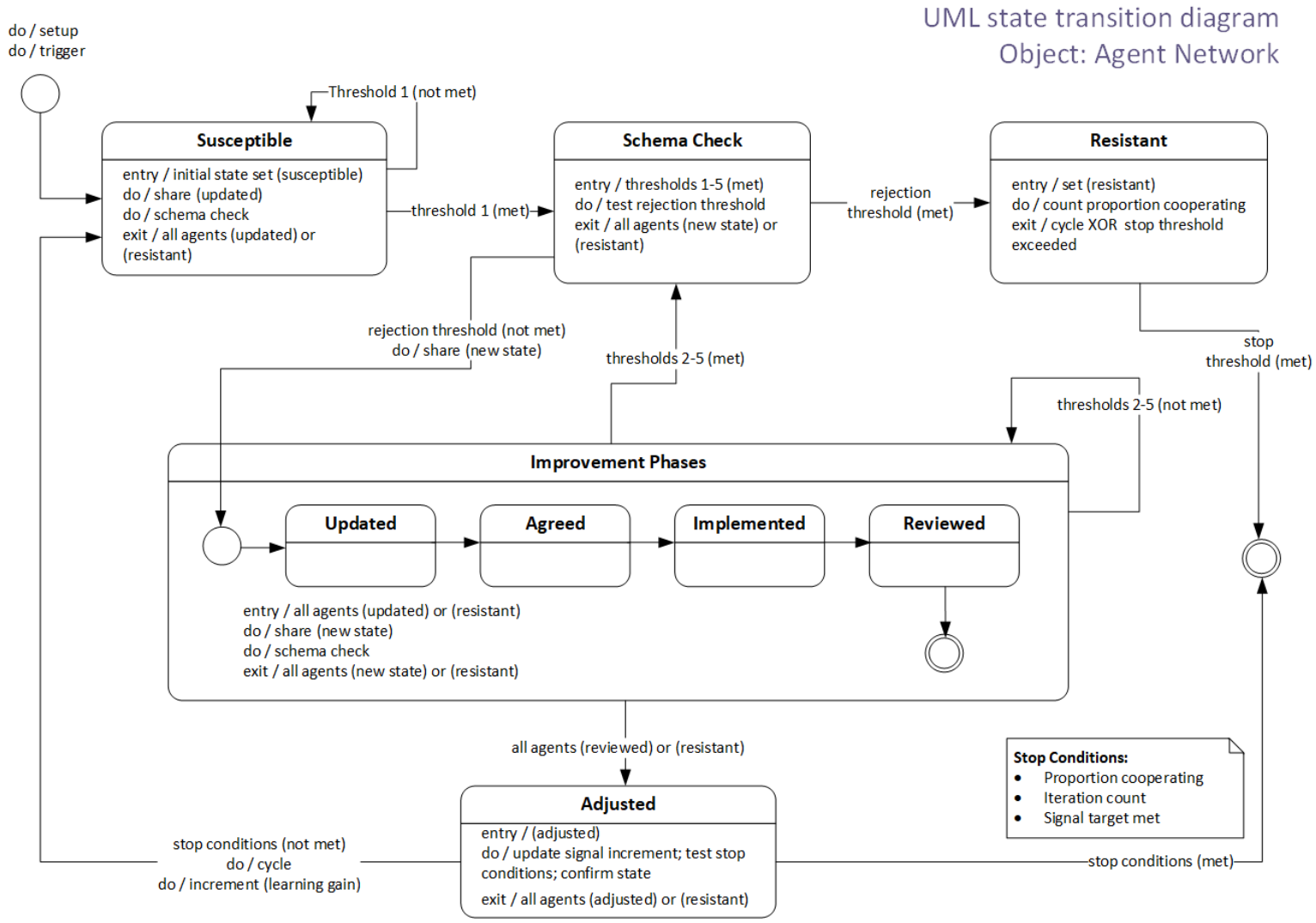
Yardley, S., Williams, H., Bowie, P., Edwards, A., Noble, S., Donaldson, L., & Carson-Stevens, A. (2022). Which human factors design issues are influencing system performance in out-of-hours community palliative care? Integration of realist approaches with an established systems analysis framework to develop mid-range programme theory. *BMJ Open, 12*(1), e048045. https://doi.org/10.1136/bmjopen-2020-048045

Yawson, R. M. (2013). Systems theory and thinking as a foundational theory in human resource development—A myth or reality? *Human Resource Development Review 12* (1), 53-85.

Young, R. A., Nelson, M., Castellon, R., & Martin, C. M. (2020). Improving quality in a complex primary care system—An example of refugee care and literature review. *Journal of Evaluation in Clinical Practice, 27*(5), 1018–1026. https://doi.org/10.1111/jep.13430

Appendices

Appendix 1: ABM State Transition Chart



Appendix 2: Secondary Data Descriptive Statistics

Distribution by (i) Healthcare context; (ii) Country; (iii) Journal

Healthcare context	Count of Articles
ARC/Older persons health	10
Primary healthcare	10
Oncology	5
Healthcare system	5
Care coordination	3
Acute: trauma care	3
Medication management	3
Coronary care	3
Emergency department	3
Professional development	3
Acute: older persons	2
Neonatal Intensive Care	2
Palliative care	2
Clinical safety	2
Acute: paediatric	2
Medical laboratory	2
Paediatric care	2
Emergency healthcare	2
ICU	2
Mental Health	2
Imaging	2
Ambulance	1
Psychiatry	1
Acute: medical	1
Medical: diabetes care	1
Surgical: hernia	1
Medical: infection prevention	1
Obstetrics and Gynaecology	1
Medical: respiratory care	1
Diagnostic procedures	1
Medical: rheumatology	1
Paediatric hospital	1
Clinical school	1
Acute: adolescent psychiatry	1
Collaborative learning	1
Hospital flow	1
Rehabilitation	1
Integrated care	1
Surgical: ORL	1
Maternal and child health	1
Medical devices	1
Grand Total	**90**

Countries	Count of Articles
United States	41
United Kingdom	11
Canada	8
Australia	4
Sweden	3
Ireland	2
United Kingdom (Scotland)	2
Netherlands	2
United Kingdom (Wales)	2
Brazil	2
Norway	2
Belgium	1
Qatar	1
Denmark	1
Saudi Arabia	1
France	1
India	1
Nigeria	1
Tanzania	1
Germany	1
Costa Rica	1
Malawi	1
Total	**90**

Journals	Count of Articles
BMC Health Services Research	9
INTERNATIONAL JOURNAL FOR QUALITY IN HEALTH CARE	5
BMJ Quality & Safety	4
BMJ Open Quality	3
Social Science & Medicine	3
Journal of Surgical Research	2
IMPLEMENTATION SCIENCE	2
Quality Management in Health Care	2
Health Policy & Planning	2
Journal of Health Organization and Management	2
International Journal of Integrated Care (IJIC)	2
JOURNAL OF INTERPROFESSIONAL CARE	1
GERIATRIC NURSING	1
Otolaryngology-Head & Neck Surgery	1
Ergonomics	1
American Journal of Medical Quality	1
EVALUATION & THE HEALTH PROFESSIONS	1
Business Process Management Journal	1

Journals	Count of Articles
BJPsych International	1
PEDIATRICS	1
American Journal of Infection Control	1
JGIM: Journal of General Internal Medicine	1
INTERNATIONAL JOURNAL OF INTEGRATED CARE	1
TOTAL QUALITY MANAGEMENT & BUSINESS EXCELLENCE	1
Journal of Clinical Nursing	1
Journal of Oncology Practice	1
Journal of Clinical Outcomes Management	1
Work	1
Journal of Evaluation in Clinical Practice	1
JONA: The Journal of Nursing Administration	1
Journal of General Internal Medicine	1
Journal of Nursing Care Quality	1
Journal of Health Management	1
Canadian Journal of Occupational Therapy	1
Journal of Health Organization & Management	1
AGING & MENTAL HEALTH	1
BMC Gastroenterology	1
Palliative Medicine	1
Journal of Infection Prevention	1
BMC Medicine	1
2019 IEEE/ACM 1st International Workshop on Software Engineering for Healthcare (SEH)	1
Applied Ergonomics	1
LEADERSHIP IN HEALTH SERVICES	1
Journal of Medical Systems	1
Nursing Administration Quarterly	1
Human Factors	1
Production Planning and Control	1
International Journal of Health Care Quality Assurance	1
BMJ Open	1
Joint Commission Journal on Quality & Patient Safety	1
RadioGraphics	1
HEALTH CARE MANAGEMENT REVIEW	1
Respiratory Care	1
Acta Paediatrica	1
J Eval Clin Pract	1
MEDICAL CARE	1
JAMA INTERNAL MEDICINE	1
Health Expectations	1
Prehospital Emergency Care	1
BMC Family Practice	1
Archives of Psychiatric Nursing	1
Family Practice	1
Advances in Neonatal Care (Lippincott Williams & Wilkins)	1
Primary Care Companion to the Journal of Clinical Psychiatry	1

Journals	Count of Articles
Children and Youth Services Review	1
Total	**90**

Appendix 3: Secondary Data References

Case ID	Reference
1	Simmons, M., Daniel, S., & Temple, M. (2019). How to manipulate friends and influence practice: Application of complexity science leads to quality improvement in laboratory sample submissions. *Journal of Infection Prevention*, *20*(2), 91–98. https://doi.org/10.1177/1757177419831348
2	Markan-Aurora, S., Miller, R., Gronemeyer, A. A., Kempf, J. S., & Hall, B. L. (2020). Implementation of a patient blood management program in a large, diverse Multi-Hospital system. *Journal of Clinical Outcomes Management*, *27*(3), 123–129. https://profiles.wustl.edu/en/publications/implementation-of-a-patient-blood-management-program-in-a-large-d
3	Storkholm, M. H., Mazzocato, P., & Savage, C. (2019). Make it complicated: a qualitative study utilizing a complexity framework to explain improvement in health care. *BMC Health Services Research*, *19*(1). https://doi.org/10.1186/s12913-019-4705-x
4	Callahan, C. M., Boustani, M., Weiner, M., Beck, R. A., Livin, L. R., Kellams, J. J., Willis, D. R., & Hendrie, H. C. (2010). Implementing dementia care models in primary care settings: The Aging Brain Care Medical Home. *Aging & Mental Health*, *15*(1), 5–12. https://doi.org/10.1080/13607861003801052
5	Anderson, R. A., Toles, M., Corazzini, K., McDaniel, R. R., & Colón-Emeric, C. (2014). Local interaction strategies and capacity for better care in nursing homes: a multiple case study. *BMC Health Services Research*, *14*(1). https://doi.org/10.1186/1472-6963-14-244
6	Young, R. A., Nelson, M., Castellon, R., & Martin, C. M. (2019). Improving quality in a complex primary care system -- an example of refugee care and literature review [Dataset]. In *Authorea*. https://doi.org/10.22541/au.157695040.00864888
7	Lennox, L., Barber, S., Stillman, N., Spitters, S., Ward, E., Marvin, V., & Reed, J. (2021). Conceptualising interventions to enhance spread in complex systems: a multisite comprehensive medication review case

	study. *BMJ Quality & Safety*, *31*(1), 31–44. https://doi.org/10.1136/bmjqs-2020-012367
8	Wilkinson, S., & Lorentzen, K. E. A. (2021). How we do it in Norway: a golden middle way for quality development of in-patient services as applied to acute adolescent psychiatry. *BJPsych International*, *19*(1), 13–15. https://doi.org/10.1192/bji.2021.7
9	Yardley, S., Williams, H., Bowie, P., Edwards, A., Noble, S., Donaldson, L., & Carson-Stevens, A. (2022). Which human factors design issues are influencing system performance in out-of-hours community palliative care? Integration of realist approaches with an established systems analysis framework to develop mid-range programme theory. *BMJ Open*, *12*(1), e048045. https://doi.org/10.1136/bmjopen-2020-048045
10	Strauss, A. T., Yeh, J., Martínez, D. A., Yenokyan, G., Yoder, J., Nehra, R., Feller, T. T., Bull-Henry, K., Stein, E. M., Hsu, L. C. H., Al-Grain, H., Zabko, C., & Fain, C. (2021). A patient-centered framework for health systems engineering in gastroenterology: improving inpatient colonoscopy bowel preparation. *BMC Gastroenterology*, *21*(1). https://doi.org/10.1186/s12876-021-01661-4
11	Rachh, P., Pendley, A. M., Duong, P. T., Hanna, T. N., & Heilbrun, M. E. (2021). Decreasing CT Acquisition Time in the Emergency Department through Lean Management Principles. *Radiographics*, *41*(3), E81–E89. https://doi.org/10.1148/rg.2021200107
12	Ramshaw, B. (2020). Applying systems and complexity science to real patient care. *Authorea (Authorea)*. https://doi.org/10.22541/au.159050270.02287037
13	Dariel, O. P. D., & Cristofalo, P. (2020). Improving patient safety in two French hospitals: why teamwork training is not enough. *Journal of Health Organisation and Management*, *34*(6), 639–653. https://doi.org/10.1108/jhom-02-2020-0045
14	Sammann, A., Chehab, L., Patel, D., Liao, J., Callcut, R. A., & Knudson, M. M. (2020). Improving efficiency and meeting expectations without compromising care on trauma surgical rounds. *Journal of Surgical Research*, *247*, 163–171. https://doi.org/10.1016/j.jss.2019.10.026

15	Bishop, S., & Waring, J. (2019). From boundary object to boundary subject; the role of the patient in coordination across complex systems of care during hospital discharge. *Social Science & Medicine*, *235*, 112370. https://doi.org/10.1016/j.socscimed.2019.112370
16	Quinn, D., Jacobson, J. O., Wong, S. L., Wollins, D. S., Gilmore, T., & Cunningham, G. (2021). Improving the Delivery of High-Quality Cancer Care in Medically Underserved Communities: a formative Evaluation method. *Quality Management in Health Care*, *30*(4), 251–258. https://doi.org/10.1097/qmh.0000000000000313
17	Pesec, M., Spigel, L., Granados, J. M. M., Bitton, A., Hirschhorn, L. R., Brizuela, J. a. J., Pignone, M., Del Rocío Sáenz, M., Schwarz, D., Del Carpio, O. V., Wilson, I. B., Méndez, E. Z., & Ratcliffe, H. L. (2021). Strengthening data collection and use for quality improvement in primary care: the case of Costa Rica. *Health Policy and Planning*, *36*(5), 740–753. https://doi.org/10.1093/heapol/czab043
18	Muttalib, F., Ballard, E., Langton, J., Malone, S., Fonseca, Y., Hansmann, A., Remy, K. E., Hovmand, P. S., & Doctor, A. (2021). Application of systems dynamics and group model building to identify barriers and facilitators to acute care delivery in a resource limited setting. *BMC Health Services Research*, *21*(1). https://doi.org/10.1186/s12913-020-06014-7
19	Alhaider, A. A., Lau, N., Davenport, P. B., & Morris, M. K. (2020). Distributed situation awareness: a health-system approach to assessing and designing patient flow management. *Ergonomics*, *63*(6), 682–709. https://doi.org/10.1080/00140139.2020.1755061
20	Burr, K., Stump, A., Bladen, R., O'Brien, P., Lemon, B., Tearl, D., Roby, M., Daley, L. M., Brown, J. R., & Hertzog, J. H. (2021). Twice-Daily Huddles improves collaborative problem solving in the respiratory care department. *Respiratory Care*, *66*(5), 822–828. https://doi.org/10.4187/respcare.07717
21	Ferreira, D. M. C., & Saurin, T. A. (2019). A complexity theory perspective of kaizen: a study in healthcare. *Production Planning & Control*, *30*(16), 1337–1353. https://doi.org/10.1080/09537287.2019.1615649

22	Brys, N., Keating, J. A., Knobloch, M. J., & Safdar, N. (2019). Engaging patients in health care epidemiology research: A case example. *American Journal of Infection Control*, *47*(2), 139–143. https://doi.org/10.1016/j.ajic.2018.08.013
23	Bergerum, C., Engström, A. K., Thor, J., & Wolmesjö, M. (2020). Patient involvement in quality improvement – a 'tug of war' or a dialogue in a learning process to improve healthcare? *BMC Health Services Research*, *20*(1). https://doi.org/10.1186/s12913-020-05970-4
24	Radcliffe, E., Kordowicz, M., Mak, C., Shefer, G., Armstrong, D., White, P., & Ashworth, M. (2020). Lean implementation within healthcare: imaging as fertile ground. *Journal of Health Organisation and Management*, *34*(8), 869–884. https://doi.org/10.1108/jhom-02-2020-0050
25	Hoyle, E., Patino, F., & Yoxall, C. W. (2019). Quality improvement programme to improve compliance with initial respiratory support guideline at preterm birth. *Acta Paediatrica*, *109*(5), 943–947. https://doi.org/10.1111/apa.15076
26	Pater, C., Sosa, T., Boyer, J., Cable, R., Egan, M. J., Knilans, T. K., Schondelmeyer, A. C., Schuler, C. L., & Madsen, N. (2020). Time series evaluation of improvement interventions to reduce alarm notifications in a paediatric hospital. *BMJ Quality & Safety*, *29*(9), 717–726. https://doi.org/10.1136/bmjqs-2019-010368
27	Yousefi-Nooraie, R., Khan, S., Gutberg, J., & Baker, G. R. (2018). A Network Analysis Perspective to Implementation: The example of health links to promote Coordinated care. *Evaluation & the Health Professions*, *42*(4), 395–421. https://doi.org/10.1177/0163278718772887
28	Cassidy, R., Tomoaia-Cotisel, A., Semwanga, A. R., Binyaruka, P., Chalabi, Z., Blanchet, K., Singh, N., Maiba, J., & Borghi, J. (2021). Understanding the maternal and child health system response to payment for performance in Tanzania using a causal loop diagram approach. *Social Science & Medicine*, *285*, 114277. https://doi.org/10.1016/j.socscimed.2021.114277
29	Vatnøy, T. K., Karlsen, T., & Dale, B. (2019). Exploring nursing competence to care for older patients in municipal in-patient acute care: A

	qualitative study. *Journal of Clinical Nursing*, *28*(17–18), 3339–3352. https://doi.org/10.1111/jocn.14914
30	Snelling, I., Benson, L., & Chambers, N. (2019). How trainee hospital doctors lead work-based projects. *Leadership in Health Services*, *33*(1), 85–100. https://doi.org/10.1108/lhs-12-2018-0064
31	McCreight, M., Gilmartin, H., Leonard, C., Mayberry, A., Kelley, L., Lippmann, B., Coy, A., Radcliff, T. A., Côté, M. J., & Burke, R. E. (2019). Practical use of process mapping to guide implementation of a care coordination program for rural veterans. *Journal of General Internal Medicine*, *34*(S1), 67–74. https://doi.org/10.1007/s11606-019-04968-x
32	Riordan, F., McGrath, N., Dinneen, S. F., Kearney, P. M., & McHugh, S. (2019). 'Sink or Swim': A qualitative study to understand how and why nurses adapt to support the implementation of integrated diabetes care. *International Journal of Integrated Care*, *19*(2), 2. https://doi.org/10.5334/ijic.4215
33	Cunney, R., Kirrane-Scott, M., Rafferty, A., Stapleton, P. J., Okafor, I., & McNamara, R. (2019). 'Start smart': using front-line ownership to improve the quality of empiric antibiotic prescribing in a paediatric hospital. *BMJ Open Quality*, *8*(3), e000445. https://doi.org/10.1136/bmjoq-2018-000445
34	T. Jachmann, "Transforming a Large Medical Organization towards Speed and Flow," 2019 IEEE/ACM 1st International Workshop on Software Engineering for Healthcare (SEH), Montreal, QC, Canada, 2019, pp. 17-20, doi: 10.1109/SEH.2019.00010.
35	Sammann, A., Chehab, L., Patel, D., & Liao, J. (2019). Why so glum? Understanding the challenges in meeting user's needs on trauma surgical rounds. *Journal of Surgical Research*, *235*, 367–372. https://doi.org/10.1016/j.jss.2018.10.006
36	Prætorius, T., Chaudhuri, A., Saddikuti, V., Hasle, P., & Singh, A. (2018). Achieving better integration in trauma care delivery in India. *Journal of Health Management*, *20*(3), 234–254. https://doi.org/10.1177/0972063418779896
37	Steele, M. L., Talley, B., & Frith, K. H. (2018). Application of the SEIPS model to analyze medication safety in a crisis residential center. *Archives*

	of Psychiatric Nursing, *32*(1), 7–11. https://doi.org/10.1016/j.apnu.2017.09.005
38	Shelby, M., Malloch, K., Shellenberger, T. D., Byrum, S. D., Mackie, L., Dilli, S., & Martin, A. D. (2019). Addressing falls among older oncology patients through Complexity science. *Nursing Administration Quarterly*, *43*(3), 280–288. https://doi.org/10.1097/naq.0000000000000360
39	Colón-Emeric, C., Corazzini, K., McConnell, E. S., Pan, W., Toles, M., Hall, R. K., Cary, M. P., Batchelor-Murphy, M., Yap, T. L., Anderson, A. L., Burd, A., Amarasekara, S., & Anderson, R. A. (2017). Effect of promoting high-quality staff interactions on fall prevention in nursing homes: a cluster-randomized trial. *JAMA Internal Medicine*, *177*(11), 1634. https://doi.org/10.1001/jamainternmed.2017.5073
40	Altman, L., Woolfenden, S., & Breen, C. (2017b). Kids GPS integrated care - leaving the Ivory Tower. *International Journal of Integrated Care*, *17*(3), 29. https://doi.org/10.5334/ijic.3141
41	White, S. A., Ewald, D., Wilson, C. M., & Rose, V. (2017). Engaging Clinicians to lead improvement of care integration. *International Journal of Integrated Care*, *17*(3), 25. https://doi.org/10.5334/ijic.3137
42	Elertson, K. M. (2017). Improving completion of the initial comprehensive health assessment for children entering foster care. *Children and Youth Services Review*, *81*, 124–128. https://doi.org/10.1016/j.childyouth.2017.08.006
43	Pearson, M., Hemsley, A., Blackwell, R. E., Pegg, L. E., & Custerson, L. (2017). Improving Hospital at Home for frail older people: insights from a quality improvement project to achieve change across regional health and social care sectors. *BMC Health Services Research*, *17*(1). https://doi.org/10.1186/s12913-017-2334-9
44	Carter, J., Dwyer, N., Roselund, J., Cote, K., Chase, P., Martorana, K., Gustafson, K. E., Bergling, E., Kamalia, R., & DeGrazia, M. (2017). Systematic changes to help parentsof medically complex infants manage medical expenses. *Advances in Neonatal Care*, *17*(6), 461–469. https://doi.org/10.1097/anc.0000000000000447

45	Halvorson, S. a. C., Wheeler, B., Willis, M., Watters, J. M., Eastman, J., O'Donnell, R., & Merkel, M. (2016). A multidisciplinary initiative to standardize intensive care to acute care transitions. *International Journal for Quality in Health Care*, *28*(5), 615–625. https://doi.org/10.1093/intqhc/mzw076
46	Kirchner, J. a. E., Woodward, E. N., Smith, J. L., Curran, G. M., Kilbourne, A. M., Owen, R. R., & Bauer, M. S. (2016). Implementation science supports core clinical competencies. *The Primary Care Companion for CNS Disorders*, *18*(6). https://doi.org/10.4088/pcc.16m02004
47	Czulada, L., Leino, P., & Willis, T. S. (2016). Partnering with a family advisor to improve communication in a pediatric intensive care unit. *American Journal of Medical Quality*, *31*(6), 520–525. https://doi.org/10.1177/1062860615600858
48	Goodridge, D., Rana, M., Harrison, E., Rotter, T., Dobson, R., Groot, G., Udod, S., & Lloyd, J. (2018). Assessing the implementation processes of a large-scale, multi-year quality improvement initiative: survey of health care providers. *BMC Health Services Research*, *18*(1). https://doi.org/10.1186/s12913-018-3045-6
49	Logan, R., Davey, P., Davie, A., Grant, S., Tully, V., Valluri, A., & Bell, S. (2018). Care bundles for acute kidney injury: a balanced accounting of the impact of implementation in an acute medical unit. *BMJ Open Quality*, *7*(4), e000392. https://doi.org/10.1136/bmjoq-2018-000392
50	Wright, R., Lowton, K., Robert, G., Grudzen, C., & Grocott, P. (2017). Emergency department staff priorities for improving palliative care provision for older people: A qualitative study. *Palliative Medicine*, *32*(2), 417–425. https://doi.org/10.1177/0269216317705789
51	Poksińska, B., & Swartling, D. (2018). From successful to sustainable Lean production – the case of a Lean Prize Award Winner. *Total Quality Management & Business Excellence*, *29*(9–10), 996–1011. https://doi.org/10.1080/14783363.2018.1486539
52	Van Harten, W. H., Goedbloed, N., Boekhout, A. H., & Heintzbergen, S. (2018). Implementing large scale fast track diagnostics in a comprehensive

	cancer center, pre- and post-measurement data. *BMC Health Services Research*, *18*(1). https://doi.org/10.1186/s12913-018-2868-5
53	McNab, D., Freestone, J., Black, C., Carson-Stevens, A., & Bowie, P. (2018). Participatory design of an improvement intervention for the primary care management of possible sepsis using the Functional Resonance Analysis Method. *BMC Medicine*, *16*(1). https://doi.org/10.1186/s12916-018-1164-x
54	Myron, R., French, C. E., Sullivan, P. A., Sathyamoorthy, G., Barlow, J., & Pomeroy, L. (2017). Professionals learning together with patients: An exploratory study of a collaborative learning Fellowship programme for healthcare improvement. *Journal of Interprofessional Care*, *32*(3), 257–265. https://doi.org/10.1080/13561820.2017.1392935
55	Rosso, C. B., & Saurin, T. A. (2018). The joint use of resilience engineering and lean production for work system design: A study in healthcare. *Applied Ergonomics*, *71*, 45–56. https://doi.org/10.1016/j.apergo.2018.04.004
56	Dang, A., Gjolaj, L. N., & Peck, H. (2018). Using rapid cycle improvement (Plan, Do, Study, Act) to design a scalable appointment scheduling system for complex oncology clinical trials at an academic cancer center. *Journal of Oncology Practice*, *14*(9), e591–e601. https://doi.org/10.1200/jop.18.00071
57	Cochran, D. S., Swartz, J. E., Elahi, B., & Smith, J. J. (2018). Using the collective system design methodology to improve a medical center emergency room performance. *Journal of Medical Systems*, *42*(12). https://doi.org/10.1007/s10916-018-1102-7
58	Elsheikh, A. M., Emam, M. S., & AlShareef, S. A. (2017). Bridging the gap between documents and practice in medication management "Documents Vitalization." *Business Process Management Journal*, *23*(4), 830–841. https://doi.org/10.1108/bpmj-02-2017-0030
59	Coleman, G., & Wiggins, L. (2017). Bringing humanity into view: action research with Qatar's ambulance service. *Journal of Health Organisation and Management*, *31*(5), 581–597. https://doi.org/10.1108/jhom-10-2016-0208

60	Eboreime, E., Nxumalo, N., Ramaswamy, R., & Eyles, J. (2018). Strengthening decentralized primary healthcare planning in Nigeria using a quality improvement model: how contexts and actors affect implementation. *Health Policy and Planning*, *33*(6), 715–728. https://doi.org/10.1093/heapol/czy042
61	Dandoy, C. E., Hausfeld, J., Flesch, L., Hawkins, D., Demmel, K., Best, D., Osterkamp, E., Bracke, T., Nagarajan, R., Jodele, S., Holt, J., Giaccone, M. J., Davies, S. M., Kotagal, U. R., & Simmons, J. M. (2015). Rapid cycle development of a multifactorial intervention achieved sustained reductions in central line-associated bloodstream infections in haematology oncology units at a children's hospital: a time series analysis. *BMJ Quality & Safety*, *25*(8), 633–643. https://doi.org/10.1136/bmjqs-2015-004450
62	Lachman, P., Linkson, L., Evans, T., Clausen, H., & Hothi, D. K. (2015). Developing person-centred analysis of harm in a paediatric hospital: a quality improvement report. *BMJ Quality & Safety*, *24*(5), 337–344. https://doi.org/10.1136/bmjqs-2014-003795
63	Hignett, S., Wolf, L., Taylor, E., & Griffiths, P. (2015). Firefighting to innovation. *Human Factors*, *57*(7), 1195–1207. https://doi.org/10.1177/0018720815593642
64	Hill, A. K., Lind, M. A., Tucker, D., Nelly, P., & Daraiseh, N. M. (2015). Measurable results: Reducing staff injuries on a specialty psychiatric unit for patients with developmental disabilities. *Work-a Journal of Prevention Assessment & Rehabilitation*, *51*(1), 99–111. https://doi.org/10.3233/wor-152014
65	Hwang, P., Hwang, D. W., & Hong, P. (2014). Lean practices for quality results: a case illustration. *International Journal of Health Care Quality Assurance*, *27*(8), 729–741. https://doi.org/10.1108/ijhcqa-03-2014-0024
66	Jiang, N., & Malkin, B. D. (2016). Use of Lean and CAHPS Surgical Care Survey to Improve Patients' Experiences with Surgical Care. *Otolaryngology-Head and Neck Surgery*, *155*(5), 743–747. https://doi.org/10.1177/0194599816657051

67	Anderson, J., Ross, A., Back, J., Duncan, M., Snell, P., Hopper, A., & Jaye, P. (2020). Beyond 'find and fix': improving quality and safety through resilient healthcare systems. *International Journal for Quality in Health Care*, *32*(3), 204–211. https://doi.org/10.1093/intqhc/mzaa007
68	Lanham, H. J., McDaniel, R. R., Crabtree, B. F., Miller, W. L., Stange, K. C., Tallia, A. F., & Nutting, P. A. (2009). How improving practice relationships among clinicians and nonclinicians can improve quality in primary Care. *The Joint Commission Journal on Quality and Patient Safety*, *35*(9), 457-AP2. https://doi.org/10.1016/s1553-7250(09)35064-3
69	Pype, P., Mertens, F., Helewaut, F., & Krystallidou, D. (2018). Healthcare teams as complex adaptive systems: understanding team behaviour through team members' perception of interpersonal interaction. *BMC Health Services Research*, *18*(1). https://doi.org/10.1186/s12913-018-3392-3
70	Clancy, T. R. (2009). Putting it altogether. *JONA: The Journal of Nursing Administration*, *39*(9), 364–367. https://doi.org/10.1097/nna.0b013e3181b3b6a8
71	Forbes-Thompson, S., Leiker, T., & Bleich, M. R. (2007). High-performing and low-performing nursing homes. *Health Care Management Review*, *32*(4), 341–351. https://doi.org/10.1097/01.hmr.0000296789.39128.f6
72	Erdek, M., & Pronovost, P. J. (2004). Improving assessment and treatment of pain in the critically ill. *International Journal for Quality in Health Care*, *16*(1), 59–64. https://doi.org/10.1093/intqhc/mzh010
73	Scott, I. A. (2004). Optimising care of acute coronary syndromes in three Australian hospitals. *International Journal for Quality in Health Care*, *16*(4), 275–284. https://doi.org/10.1093/intqhc/mzh051
74	Margolis, P. A., Stevens, R. B., Bordley, W. C., Stuart, J., Harlan, C., Keyes-Elstein, L., & Wisseh, S. (2001). From Concept to Application: The impact of a Community-Wide intervention to improve the delivery of preventive services to children. *Pediatrics*, *108*(3), e42. https://doi.org/10.1542/peds.108.3.e42
75	Bonzo, S. M., & Liker, J. K. (2016). The impact of Cross-Unit coordination and process improvement efforts in Anatomic Pathology: A

	Systems Perspective. *Quality Management in Health Care*, *25*(3), 141–148. https://doi.org/10.1097/qmh.0000000000000103
76	Ngo, J., Lau, D., Ploquin, J., Receveur, T., Stassen, K., & Del Castilho, C. (2022). Improving incident reporting among physicians at south health campus hospital. *BMJ Open Quality*, *11*(4), e001945. https://doi.org/10.1136/bmjoq-2022-001945
77	Goderis, G., Borgermans, L., Mathieu, C., Van Den Broeke, C., Hannes, K., Heyrman, J., & Grol, R. (2009). Barriers and facilitators to evidence based care of type 2 diabetes patients: experiences of general practitioners participating to a quality improvement program. *Implementation Science*, *4*(1). https://doi.org/10.1186/1748-5908-4-41
78	Rantz, M., Flesner, M., & Zwygart-Stauffacher, M. (2010). Improving care in Nursing Homes using Quality Measures/Indicators and Complexity Science. *Journal of Nursing Care Quality*, *25*(1), 5–12. https://doi.org/10.1097/ncq.0b013e3181c12b0f
79	Kauth, M. R., Sullivan, G., Blevins, D. G., Cully, J. A., Landes, R. D., Said, Q., & Teasdale, T. A. (2010). Employing external facilitation to implement cognitive behavioral therapy in VA clinics: a pilot study. *Implementation Science*, *5*(1). https://doi.org/10.1186/1748-5908-5-75
80	Crabtree, B. F., Nutting, P. A., Miller, W. L., McDaniel, R. R., Stange, K. C., Jaén, C. R., & Stewart, E. (2011). Primary Care Practice Transformation is Hard Work. *Medical Care*, *49*, S28–S35. https://doi.org/10.1097/mlr.0b013e3181cad65c
81	Murray, M., Higuchi, K. S., Edwards, N., Greenough, M., & Hoogeveen, K. (2011). Practice change in Long-Term Care Facilities in Ontario: Views of managers and staff. *Geriatric Nursing*, *32*(6), 418–428. https://doi.org/10.1016/j.gerinurse.2011.08.001
82	Rathert, C., Brandt, J., & Williams, E. S. (2011). Putting the 'patient' in patient safety: a qualitative study of consumer experiences. *Health Expectations*, *15*(3), 327–336. https://doi.org/10.1111/j.1369-7625.2011.00685.x
83	Villeneuve, M., & Shulha, L. M. (2012). Learning together for effective collaboration in school-based occupational therapy practice. *Canadian*

	Journal of Occupational Therapy, *79*(5), 293–302. https://doi.org/10.2182/cjot.2012.79.5.5
84	Booth, B., Zwar, N., & Harris, M. (2013). Healthcare improvement as planned system change or complex responsive processes? a longitudinal case study in general practice. *BMC Family Practice*, *14*(1). https://doi.org/10.1186/1471-2296-14-51
85	Cranley, L., Hoben, M., Yeung, J., Estabrooks, C. A., Norton, P., & Wagg, A. (2018). SCOPEOUT: sustainability and spread of quality improvement activities in long-term care- a mixed methods approach. *BMC Health Services Research*, *18*(1). https://doi.org/10.1186/s12913-018-2978-0
86	Bidassie, B., Davies, M. L., Stark, R., & Boushon, B. (2014). VA experience in implementing Patient-Centered Medical Home using a Breakthrough Series collaborative. *Journal of General Internal Medicine*, *29*(S2), 563–571. https://doi.org/10.1007/s11606-014-2773-5
87	Hilts, L., Howard, M., Price, D., Risdon, C., Agarwal, G., & Childs, A. M. (2012). Helping primary care teams emerge through a quality improvement program. *Family Practice*, *30*(2), 204–211. https://doi.org/10.1093/fampra/cms056
88	Essén, A., & Lindblad, S. (2013). Innovation as emergence in healthcare: Unpacking change from within. *Social Science & Medicine*, *93*, 203–211. https://doi.org/10.1016/j.socscimed.2012.08.035
89	Talbert-Slagle, K., Berg, D. N., & Bradley, E. H. (2013). Innovation spread: lessons from HIV. *International Journal for Quality in Health Care*, *25*(4), 352–356. https://doi.org/10.1093/intqhc/mzt038
90	Chen, J., Awasthi, A., Shechter, S. M., Atkins, D., Lemke, L., Fisher, L. D., & Dodek, P. (2013). Using operations research to plan improvement of the transport of critically ill patients. *Prehospital Emergency Care*, *17*(4), 466–474. https://doi.org/10.3109/10903127.2013.811561

Appendix 4: Thematic Coding Examples

CQIN construct	Coding example 1	Case ID	Coding example 2	Case ID	Coding example 3	Case ID
Agent network	*Through the use of a multidisciplinary improvement team approach including a family advisor, the number of documented communication events between a patient/family and the nurses and physicians significantly increased and has been sustained for more than 2 years.*	[47]	*Care coordinators are responsible for facilitating shared care models, identifying opportunities to localise care and guiding the child and their family in navigating the health system.*	[41]	*Forming networks, cross-functional educational days, re-organising operational teams influenced the patterns of relationships and thereby caused people to think differently an operate in new constellations.*	[52]
Schema share effectiveness	*Our use of PDSA cycles demonstrated the significance of working collaboratively across teams. Building productive working relationships between teams required recognising strengths and weaknesses and working together to reconfigure services in a way that was patient-centered.*	[43]	*The aim is for everyone to understand their role in a child's care plan.*	[41]	*A one-day summit was held involving approximately 40 key stakeholders including physicians, house officers, nursing staff, and pharmacists, as well as representatives from transportation services, environmental services and bed flow management.*	[46]

CQIN construct	Coding example 1	Case ID	Coding example 2	Case ID	Coding example 3	Case ID
Problem solving effectiveness	*Participants mapped the existing process of patient transfer from ICU to ACU using a swim-lane flowchart… Lean methodology was employed to perform an analysis of waste and opportunities for standardization.*	[45]	*The following key concepts and interventions are believed to have led to the measurable reduction in staffing injuries: (1) Safety team consisting of diverse membership with measurable goals, strategy map (Key Driver Diagram) to guide the work, and 90 day cycles for strategy evaluation; (2) Weekly focus on performance via run charts .... (3) Plan, Do, Study, Act rapid intervention testing process for universal precautions, initial assessments, protective equipment decision key, personal protective equipment guidelines, leadership rounds, huddles, and patient identification board; and (4) critical incident reviews.*	[65]	*This case illustrates that it may be possible to become more successful if the quality improvement approach combines QI methods with an understanding of the importance for learning and shared responsibility to deal with co-existing degrees of contextual complexity in modern healthcare.*	[3]
System signals	*Email communication on comprehensive health assessment rates was sent to case managers each week.*	[42]	*In summary, data and feedback processes were described by some participants as a means to initiate and implement change.*	[82]	*The introduction of a simple, real-time bedside reporting tool and risk management process facilitated active engagement with families and patients, and the opportunity for disclosure and learning for staff.*	[63]

CQIN construct	Coding example 1	Case ID	Coding example 2	Case ID	Coding example 3	Case ID
Coevolution constraints	*Thus, political actors at this local level tend to prioritize initiatives with adequate political capital or incentives, as against improvement of socio-economic indices of the populace.*	[60]	*External factors influence the collaboration. When the GP is less available (e.g. due to workload in general practice in winter time), the [palliative home care team] nurse takes over the coordinating role from the GP.*	[69]	*Coevolution is an important complexity element in the story. Key aspects of good chronic illness care developed as a consequence of efforts in unrelated area - both clinical and business.*	[84]
Autonomy	*Clinicians know what is wrong with the system and are well placed to test solutions if allowed to do so.*	[41]	*If all document vitalization successes were [attributed] to one factor, it would be "empowerment". Once leaders have the courage to listen to frontline staff voice and allow them to do things differently, their staff will surprise their organizations with the marvels of their creations.*	[58]	*Professional freedom and medical professionalism are further deeply rooted values that persisted throughout the process. There was never any doubt regarding the fact that the rheumatologist association……should determine the content and functionality of the innovation.*	[88]
Focus / Time	*The dynamic of the practice meeting revealed the pattern of organizing that fostered change and improvement over the years. Matters were discussed according to an agenda that began to be defined when the practice meeting was scheduled.*	[84]	*Managers continually encouraged honest and open dialog in their interactions with staff - key to effective leadership in complex situations as they help generate a diversity of ideas for further probing.*	[3]	*More than 75% of respondents indicated that neither sufficient training nor resources (collective action) had been made available to them for implementation of Lean.*	[48]

CQIN construct	Coding example 1	Case ID	Coding example 2	Case ID	Coding example 3	Case ID
Learning gain	*We found that rapid cycles of PDSA tests facilitated iterative learning that in turn enabled service change to occur at pace.*	[43]	*We implemented several consecutive interventions to improve pain assessment.*	[72]	*Learning is also a social act. One way to improve healthcare quality is to encourage a culture of learning - learning from mistakes, learning by doing, and learning by experiencing history richly.*	[68]
Problem space	*Time was identified as a critical element in standardizing expectations and predictability of the transfer process. After reviewing baseline data, summit participants determined that a transfer should be completed within 120 min.*	[45]	*We viewed application of CBT training and development of skill competency as complex, developmental tasks that would require the facilitator to employ a range of enabling strategies varying with the therapist's self-efficacy, skill competency, and situation.*	[79]	*Initially they sensed that there was a possibility for improvement. This guided their change strategy of extensive analysis in interdisciplinary camps, which made it possible to probe for new responses.*	[3]
Interpersonal factors	*Improved multidisciplinary teamwork across sectors [….] maturing of organisational relationships.*	[41]	*The fellowship is modelled on collaborative learning theory, which suggests that social interaction is an integral part of learning and multiple perspectives strengthen this process.*	[54]	*However, in the single-minded pursuit of improvement science, we suggest that something gets lost - the appreciation of humanity, the art of conversation, the ability of leaders and staff to listen to each other, to explore their experience, to voice their concerns and hope.*	[59]

CQIN construct	Coding example 1	Case ID	Coding example 2	Case ID	Coding example 3	Case ID
Option generation	*During Cycle 2, a new system for [health benefits coordinator] assessment was developed and implemented. A bedside nurse-initiated screening tool was developed that would rapidly trigger an evaluation by the [social worker], nurse case manager and health benefits coordinator.*	[44]	*Interventions [included] Clinical guidelines, reminder tools and educational interventions; 6-monthly performance feedback; pharmacist-mediated patient education program; and facilitation of multidisciplinary review of practices.*	[73]	*Ensuring and making clear the alignment between principles and purpose of the gesture, and its inherent power (such as financial incentives, regulations or sanctions), may influence the responses from which will emerge changing patterns of organizing.*	[84]
Budget	*To prevent blockage based on capacity worries, 4 [hours] of additional MRI-capacity per week had to be decided upon as some senior clinicians were convinced that some degree of redundancy would prove to be an essential success factor.*	[52]	*Government, particularly at state and [local government area] levels, needs to show leadership, technical political and sustainable financial commitment.*	[60]	*Similarly, financial viability was identified as a need, particularly for [practice doctor 3], and that stimulated clear planned improvement activity.*	[84]

Appendix 5: Full Secondary Data Construct-Case Matrix

1=present/observed; 0=not present/not observed

Case ID	Schema share effectiveness	Agent network	Problem solving effectiveness	System signals	Autonomy	Learning gain	Coevolution constraints	Focus / Time	Problem space	Option generation	Interpersonal factors	Budget	QCA 1/0	Count
26	1	1	1	1	1	1	1	1	1	0	0	1	1	10
7	1	1	1	1	1	0	1	1	0	1	1	0	1	9
14	0	1	1	1	1	1	1	1	1	1	0	0	1	9
8	1	1	1	1	1	1	0	1	0	0	1	0	1	8
16	1	1	1	0	1	1	1	1	0	0	0	1	1	8
20	1	1	1	0	1	1	1	1	0	1	0	0	1	8
24	1	1	1	1	1	0	0	1	0	1	1	0	1	8
25	0	1	1	1	1	1	1	1	0	1	0	0	1	8
70	1	1	1	1	0	1	1	0	1	1	0	0	1	8
3	1	1	1	0	0	1	0	1	1	1	0	0	1	7
4	1	1	1	0	1	0	0	1	0	0	1	1	1	7
38	1	1	1	1	1	0	0	0	1	1	0	0	1	7
41	1	1	1	1	1	1	0	0	0	0	1	0	1	7
51	1	1	1	1	1	0	0	0	0	1	0	1	1	7
86	1	0	1	1	0	1	1	0	1	0	0	1	1	7
9	1	1	1	0	0	1	1	0	0	0	1	0	1	6
13	1	1	0	0	0	0	1	1	0	0	1	1	1	6
18	1	0	1	0	0	0	1	1	1	1	0	0	1	6
22	1	1	1	0	1	0	0	0	1	0	0	1	1	6
23	1	1	0	0	1	1	0	1	0	0	1	0	1	6
31	1	1	1	0	0	1	1	0	1	0	0	0	1	6
32	1	1	1	0	1	0	1	0	0	0	1	0	1	6
40	1	1	1	1	0	1	1	0	0	0	0	0	1	6
45	1	1	1	0	1	0	0	1	1	0	0	0	1	6
52	0	0	1	0	1	1	0	0	1	0	1	1	1	6
64	1	0	1	1	1	0	0	1	1	0	0	0	1	6
65	1	1	1	0	1	0	0	1	0	0	0	1	1	6
71	1	1	0	1	1	0	0	0	0	1	1	0	1	6
73	1	1	0	1	1	0	0	0	1	1	0	0	1	6
85	1	1	1	0	0	0	0	1	0	0	1	1	1	6
87	1	1	1	0	0	0	1	1	0	0	1	0	1	6
88	1	1	0	1	1	1	0	0	0	1	0	0	1	6
5	1	0	1	1	0	0	0	0	1	1	0	0	1	5
6	0	1	0	1	1	0	1	0	0	0	0	1	1	5
19	1	1	0	1	0	0	0	1	1	0	0	0	1	5
21	0	0	1	1	1	0	1	0	1	0	0	0	1	5
33	0	1	1	1	1	1	0	0	0	0	0	0	1	5
43	1	0	1	0	0	1	1	0	1	0	0	0	1	5
46	1	0	0	1	0	1	0	0	1	1	0	0	1	5

Case ID	Schema share effectiveness	Agent network	Problem solving effectiveness	System signals	Autonomy	Learning gain	Coevolution constraints	Focus / Time	Problem space	Option generation	Interpersonal factors	Budget	QCA 1/0	Count
47	0	1	1	0	1	0	0	1	1	0	0	0	1	5
53	0	1	0	1	1	0	0	0	1	1	0	0	1	5
54	1	1	0	0	0	1	0	0	0	1	1	0	1	5
61	0	0	1	1	1	1	0	0	1	0	0	0	1	5
62	1	0	1	1	0	1	0	0	0	1	0	0	1	5
68	1	1	0	1	0	1	0	0	0	0	1	0	1	5
75	1	1	1	1	0	0	1	0	0	0	0	0	1	5
80	1	1	0	0	0	1	1	0	0	0	1	0	1	5
83	1	1	0	0	0	1	0	1	1	0	0	0	1	5
84	1	1	0	0	0	0	1	1	0	0	0	1	1	5
89	1	1	0	0	1	0	1	0	0	1	0	0	1	5
11	0	1	1	1	0	0	0	1	0	0	0	0	1	4
12	0	1	1	1	0	1	0	0	0	0	0	0	1	4
15	1	1	0	1	0	0	1	0	0	0	0	0	1	4
27	1	1	0	0	0	0	1	0	0	0	1	0	1	4
37	1	0	1	1	0	0	0	0	0	1	0	0	1	4
44	0	1	1	0	0	1	0	0	0	1	0	0	1	4
50	1	1	0	1	0	0	0	0	1	0	0	0	1	4
56	0	1	1	1	0	0	0	0	1	0	0	0	1	4
57	1	1	1	0	0	0	0	0	1	0	0	0	1	4
67	1	1	1	0	0	0	0	0	0	0	1	0	1	4
69	1	1	0	0	0	0	1	0	0	0	1	0	1	4
72	0	0	1	1	0	1	0	0	0	1	0	0	1	4
76	1	1	0	1	0	0	0	0	1	0	0	0	1	4
77	1	0	1	1	0	0	0	0	0	0	1	0	1	4
78	1	0	1	1	0	0	0	1	0	0	0	0	1	4
1	1	0	0	1	0	1	0	0	0	0	0	0	1	3
2	1	1	0	1	0	0	0	0	0	0	0	0	1	3
10	0	1	1	1	0	0	0	0	0	0	0	0	1	3
29	1	0	1	0	0	1	0	0	0	0	0	0	1	3
30	1	0	1	0	0	0	0	0	0	0	1	0	1	3
34	0	0	1	1	0	0	1	0	0	0	0	0	1	3
35	1	1	1	0	0	0	0	0	0	0	0	0	1	3
58	0	0	1	0	1	0	0	0	0	1	0	0	1	3
59	1	1	0	0	0	0	0	0	0	0	1	0	1	3
60	0	1	0	0	0	0	1	0	0	0	0	1	1	3
74	1	1	0	0	0	0	1	0	0	0	0	0	1	3
79	1	0	1	0	0	0	0	0	1	0	0	0	1	3
81	1	0	1	1	0	0	0	0	0	0	0	0	1	3
82	0	1	0	1	0	0	0	1	0	0	0	0	1	3
90	0	0	1	0	0	0	0	0	1	1	0	0	1	3
17	1	0	0	1	0	0	0	0	0	0	0	0	1	2

Case ID	Schema share effectiveness	Agent network	Problem solving effectiveness	System signals	Autonomy	Learning gain	Coevolution constraints	Focus / Time	Problem space	Option generation	Interpersonal factors	Budget	QCA 1/0	Count
28	0	0	0	1	0	0	1	0	0	0	0	0	1	2
36	0	1	0	0	0	0	0	0	1	0	0	0	1	2
42	1	0	0	1	0	0	0	0	0	0	0	0	1	2
55	0	0	1	0	0	0	0	0	1	0	0	0	1	2
63	0	0	1	0	0	0	0	0	1	0	0	0	1	2
66	0	0	1	0	0	1	0	0	0	0	0	0	1	2
39	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	62	59	57	46	30	31	29	25	31	24	22	13	Sum	
	0.69	0.66	0.63	0.51	0.33	0.34	0.32	0.28	0.34	0.27	0.24	0.14	%	

Appendix 6: Applied Case Study Survey

	Meta data
	Response ID
	Start time
	Completion time
	Email (optional)
	Name (optional)

	Consent
Q1	I understand that taking part in this study is voluntary (my choice) and that I may withdraw from the study at any time
Q2	I have read, or have had read to me in my first language, the participant Information above
Q3	I have had sufficient time to consider whether or not to participate in this study
Q4	I wish to receive a summary of the results from the study
Q5	Contact email to receive study summary (optional)

	System Signals
Q6	Please list the improvement outcome(s) for the antipsychotic deprescribing project that you are aware of
Q7	Improvement outcome feedback strength (Ranked scale 1-5)
Q8	How are you made aware of improvement outcomes related to the antipsychotic deprescribing project?
Q9	Other system feedback strength (Ranked scale 1-5)
Q10	How are you made aware of other system feedback related to the antipsychotic deprescribing project?

	Agent Network
Q11	Are the members of the Medications Advisory Committee part of the core improvement network or part of a related network?
Q12	Comments (optional)
Q13	Are Prescribers part of the core improvement network or part of a related network?
Q14	Comments (optional)
Q15	Are ARC facilities clinical leads part of the core improvement network or part of a related network?
Q16	Comments (optional)
Q17	Are ARC facilities clinical staff part of the core improvement network or part of a related network?
Q18	Comments (optional)
Q19	Are ARC facilities management part of the core improvement network or part of a related network?
Q20	Comments (optional)
Q21	Are Patients part of the core improvement network or part of a related network?
Q22	Comments (optional)
Q23	Are Patient's families/whanau part of the core improvement network or part of a related network?
Q24	Comments (optional)
Q25	If you think there are additional agent groups not included above, please list them here and indicate whether you think they are part of your network or a related network.
Q26	If you have experienced any difficulties working with some of the agent groups, please briefly list these difficulties here.

	Schema Share Effectiveness and Problem Solving Effectiveness
Q27	Within the context of the antipsychotic deprescribing project, how do you know whether your collective problem solving actions are working?
Q28	Team alignment of project goals (Ranked scale 1-5)
Q29	Please describe the improvement goals as you understand them
Q30	Team alignment of project interventions (Ranked scale 1-5)
Q31	What do you think is the most important improvement action to take at this stage of your project?
Q32	What do you consider to be the major constraint or obstacle to the project success?
Q33	Please describe the way your network tries to solve problems. You can use a formal description such as 'plan-do-study-act' or simply describe what happens
Q34	Problem solving effectiveness (Ranked scale 1-5)
	Success Prediction
Q35	Success prediction (Ranked scale 1-5)

Appendix 7: Research Impact Progress Note

In early 2024 the author was contacted by two researchers from the James M. Anderson Center for Health Systems Excellence at Cincinnati Children's Hospital (CCH). The CQIN agent-based model, made available on the NetLogo modelling commons, had caught their attention. Professor Michael Seid heads a research lab exploring the science of collaborative health networks, using a range of methods including agent-based modelling. In the view of Professor Seid, modelling QI is still an under-researched field with ample scope to share ideas and work cooperatively on common research goals. At the time of submission of this thesis, we are exploring several collaboration avenues: (i) using network maturity survey data from CCH learning networks to calibrate the CQIN Bayesian network risk assessment model; (ii) accessing wider expert input, also to calibrate CQIN alongside data, and (iii) constructing independent agent-based models built on common principles to compare the results and assess the concurrent validity of simulation findings via independent models.