Categories of context in realist evaluation

Abstract

Realist evaluation has become widespread partly because of its sensitivity to the influence of contexts upon policy implementation. In many such evaluations the range of contexts considered relevant nevertheless remains disparate and under-conceptualised. This paper uses findings from a realist evaluation of English Patient Safety Collaboratives (PSC) during 2015-18 to develop a realist taxonomy of contexts, differentiating contexts according to how they affect the corresponding policy mechanism. By analysing the main context-mechanism-outcome configurations that made up the English PSCs we derive a taxonomy of the contexts that affected the implementation and outcomes of the PSCs. The categories of context were: structural (network, hierarchy, market and organisation contexts); resource-based (actors, material, financial); motivational (receptivity, outcome headroom) and temporal (continuity, history and convergence). To the categories found in previous studies this study adds the three temporal contexts.

Keywords

realist evaluation; context; patient safety collaborative; time; NHS

Categories of context in realist evaluation

Realist evaluation has become widely-used partly because of its sensitivity towards the implications of the contexts of policy implementation. Nevertheless, the range of contexts considered relevant is disparate and often under-conceptualised. Health researchers, for example, have attempted to list comprehensively the factors affecting programme implementation (Kaplan et al. 2012; Damschroder and Lowery 2013; Vanderkruik and McPherson 2017), but without exploiting realist methodology. All but one of Rog's (2012) categories of context are contexts of evaluation work, not of the mechanisms being evaluated. Even some putatively realist studies either ignore context or treat it as a residual explanatory dumping-ground, categorising everything which was neither mechanism nor outcome as 'context'. Patient Safety Collaboratives (PSC) are widely used to increase patient safety but their outcomes appear to be heavily context-dependent. The purpose and contribution of this paper is therefore to use findings from a realist evaluation of English PSCs to start developing a realist taxonomy of contexts, one that is theoretically-based in differentiating contexts according to how they alter the workings of the focal policy mechanism. Previous studies suggest an initial taxonomy distinguishing motivational contexts (receptivity; outcome headroom); resource contexts (finance; sufficient actors; material input availability), structural (network; market; hierarchical; organisational) contexts. Our findings suggest adding temporal (time-dependent) contexts: continuity; history; and convergence.

1   'Context' in realist evaluation

To be useful for decision-makers evaluations need to identify what policy interventions work, in which context, and for whom, rather than merely 'does it work?'. A premiss of realist evaluation is that every policy implicitly assumes that activating specific policy mechanisms will, in a suitable

context, produce pre-specified policy outcomes. The study of generative mechanisms is what makes such evaluations realist. The evaluative aspect is to assess empirically the success or failure of a policy *in the policy-makers' own terms.* That requires clearly distinguishing intended from actual (observed) outcomes and, since policies tend to be contested, the outcomes that policy-makers wanted from those that the implementers, policy subjects (those whose behaviour the policy aims to change) or others wanted.

The definitional boundary between mechanisms and context is often unclear, even disputed (Marchal et al. 2012; Astbury and Leeuw 2010; Rog 2012). The same resources, information and knowledge may function as mechanism for one policy and context for another (Coldwell 2019). According to circumstances, a mechanism may be a social (policy, managerial, organisational) or a physical (clinical, technical) generative structure (Dixon-Woods 2014), or a combination. Always, though, a policy implicitly rests on a prospective programme theory. It assumes – possibly wrongly – that a set of policy subjects will respond in foreseen ways to new forms of persuasion, opportunities or incentives (including affective incentives such as appeals to identity) (Pawson and Tilley 1997). A programme mechanism is defined, by realism, as the complex of action, processes and generative structures through which policy-makers try to get the policy subjects to respond accordingly, thereby creating the outcomes that the policy-maker desires (The RAMESES II Project 2017a). Those implementing a policy may include more actors than those whom the policy-makers hold formally responsible (e.g. when relying on mass media naming-and-shaming offenders (Pawson 2008)). Some mechanisms (e.g. changing prescribing behaviour) take effect through individual, others (e.g. teamwork changes) through collective, action. Mechanisms have the characteristics that Bradford-Hill (1965) first described: strong association, specificity, consistency, experimental manipulation, succession, dose-response gradient, and coherence with other known mechanisms (Lewis et al. 2020). They usually decompose into constituent, nested mediating sub-

mechanisms (Dixon-Woods 2014) (not to be confused with the whole mechanism (Lewis et al. 2020) nor mediators in a statistical sense). There may be multiple, even alternative or conflicting, parallel mechanisms which together increase the chance of producing the intended outcomes (Aveling, Parker, and Dixon-Woods 2016; Manley et al. 2017) but are also harder to implement than simple mechanisms (Herepath, Kitchener, and Waring 2015; Kaplan et al. 2012; Knight et al. 2012; J. C. Øvretveit et al. 2011). Under-defined policy mechanisms may also give the implementers some discretion (Bastien 2009). The mechanism by which a policy putatively works is not necessarily the same thing as the preliminary activities which set it in motion (Astbury and Leeuw 2010).

In general, a context is therefore defined relative to the mechanism (policy or practice intervention) that it is the context of. Realism defines contexts as conditions which increase or decrease the outcomes that the mechanism produces (The RAMESES II Project 2017b). They cannot by themselves produce those outcomes (Lewis et al. 2020) but they do restrict (Bergen and White 2005), deflect (Jones 2007) or amplify (Sheaff 1997) that mechanism's own capacity to produce the outcomes (Pawson 2008). That is because they are 'dynamic, agentic, relational [i.e. relate to the mechanism], historically located, immanent and complex' (Coldwell 2019, p.99). They include conditions both external and internal to the group(s) who implement the mechanism (Vanderkruik and McPherson 2017; Kaplan et al. 2012). Although these conditions may pre-exist the focal mechanism, they only function as its context when and where attempts are made to implement it (Coldwell 2019). Which contextual factors are most influential may change as a project moves through design, implementation, scale-up, spread, and sustainability stages (Vanderkruik and McPherson 2017) and a mechanism may re-shape its own context (Barnes, Matka, and Sullivan 2003). Because contexts in a realist sense have traceable connections with the focal mechanism and causally affect its outcomes they differ from mere settings, which do not.

As with the concept of mechanism (Astbury and Leeuw 2010), an empirical starting-point for

more sharply conceptualising context in a realist way might therefore be a taxonomy of the kinds of contexts that the innumerable realist evaluations report, but then two complications arise. First, realist evaluation can and often does re-use, indeed build upon, findings from studies which did not use realist methods or frameworks but nevertheless report what realists would recognise as a policy's contexts, mechanisms and outcomes (Pawson et al. 2005). So interpreted, findings from these studies can also contribute to a realist conceptualisation of what kinds of context affect the operation of generative mechanisms. But that exacerbates the second complication, the sheer volume of material. One can temporarily mitigate that problem by starting from one policy area to formulate an initial taxonomy which can later be tested and nuanced, perhaps also enlarged, by trying to apply it more widely.

As an evidence base from which to develop a realist taxonomy of contexts we therefore use a realist evaluation of PSCs in England. For present purposes this starting-point has several advantages. Besides these new data there are many evaluations of other Collaboratives to draw upon, some of them realist. Many report what contexts make Collaboratives succeed or fail. Collaboratives are reported in (at least) Australia, Canada (Rossiter et al. 2017), France (Anon. 2013), Mexico (Barceló et al. 2010), the Netherlands (Dückers et al. 2009; Broer, Nieboer, and Bal 2012), Norway, Peru, Russia, Sweden (Algurén et al. 2019), the UK, USA and several African countries (Nembhard 2012), making this starting-point internationally relevant. Following avoidable quality and safety failures at Mid-Staffordshire Hospital (The Mid Staffordshire NHS Foundation Trust Inquiry 2010) the UK government mandated the NHS to establish, among other things, Patient Safety Collaboratives (Department of Health 2014). They were regional networks of clinical teams, each of which developed practical methods for improving patient safety in its own workplace. The teams collaboratively compared activities, methods and results in order to make their next iterations of these activities more effective (see

https://improvement.nhs.uk/resources/patient-safety-improvement-programmes/). The policy drew mainly upon three US prototypes: the New England Quality Improvement Collaboratives; the Institute for Clinical Systems Improvement approach of developing practice guidelines and promoting culture change across healthcare providers (Hatahet, Bowhan, and Clough 2004; Solberg 2005); and the Institute of Healthcare Improvement 'Breakthrough Series' Quality Improvement Collaboratives approach of periodically reporting and critically comparing mortality, morbidity and medical practice across its member-organisations (Hulscher, Schouten, and Grol 2009). We use the term 'PSC' for the English NHS collaboratives specifically and 'Collaboratives' for other collaboratives or collaboratives generally. Without using the phrase 'context-mechanism-outcome configuration' (CMOC), the main English main policy statements (Berwick 2013; Keogh 2013; Department of Health 2014) nevertheless stated the mechanisms and contexts through which PCS would work: the policy-makers' prospective programme theory (Figure 1).

[FIGURE 1 HERE]

2 The Collaborative Mechanism

The eponymous collaboration is between (and within) multi-professional clinical teams in different providers who compare problems, possible solutions or obstacles to them, implementation methods and the outcomes of the attempted solutions. A network coordinated by a participating provider or a separate body shares this learning between the clinical teams (Curry et al. 2017; Dückers et al. 2009; Kilo 1998; Manser 2009; Shaw et al. 2012; Singer and Vogus 2013), provides clinical expertise (Herepath, Kitchener, and Waring 2015), pools teams' data about their quality improvement (QI) activities and encourages mutual emulation. Such activities, it is assumed,

strengthen safety culture within the participating organisations (Benn et al. 2012; Brewer 2006; Curry et al. 2017; Dillon et al. 2016; Ferlie and Shortell 2001; Kaplan et al. 2010; Krein et al. 2010; Mardon et al. 2010; Parand et al. 2012), motivating clinical teams to:

1. Discover and apply formal evidence, and consult experts, so as to make their clinical work-practices and care safer.

2. Learn and apply QI techniques, e.g. how to measure clinical quality and safety, how to use repeated measurements, and how to counteract the cognitive and motivational 'human factors' barriers (Rasmussen and Pedersen 1984) to recognising and remedying quality and safety defects in healthcare.

3. Apply this knowledge through plan-do-study-act (PDSA) cycles (of which many variants and names exist) of assessing the current quality or safety of some aspect of clinical work, measuring it, implementing an attempt to improve it, re-measuring and revising the changed work-process again as necessary.

These activities usually focus on a specific problem or activity (e.g. preventing infection during or after surgery). They are a Collaborative's core mechanisms.

Extensive research, including systematic reviews (Cunningham et al. 2011; Franco and Marquez 2011; Hulscher, Schouten, and Grol 2009; Kringos et al. 2015; Nadeem et al. 2013; J. Øvretveit et al. 2002; Wells et al. 2018), gives mixed findings about Collaboratives' outcomes. The evidence that Collaboratives improve healthcare quality and safety is 'positive but limited' (Schouten et al. 2008). An early review found that around 30% of health providers involved in collaboratives improved patient safety, 40% achieved little and 30% dropped out (Øvretveit et al. 2002). A later review reported below 20% drop-out and that the reported quality and safety gains 'ranged from modest to very impressive, with a median of 12% improvement (range 4%–61%)' (Wells et al. 2018, p.233). Another (Nadeem et al. 2013) found that between 73% and 83% of

studies reported that at least one primary effect measure had improved, depending how strictly one defined 'improve'. That review and another (Davies et al. 2008) reported mixed results with respect to patient experiences. An overview of 27 Collaboratives in twelve middle- and lower-income countries (Franco and Marquez 2011) found that 87% of the Collaboratives showed 'performance' levels increases of 80%. The most recent systematic review (Wells et al. 2018) included primary studies up to 2014. Subsequent primary studies (until early 2020) still report mixed findings. Some (Curry et al. 2017; Fulton et al. 2017; Smith et al. 2017) continue to endorse PSC effectiveness. Despite any publication bias against reporting them, reports of apparently ineffective Collaboratives also continue to appear (e.g. Herepath, Kitchener, and Waring 2015; Russ et al. 2015; Algurén et al. 2019). Attempts to replicate the Michigan Keystone Collaborative had at best equivocal effects on clinical practice in the UK (Dixon-Woods, Leslie, et al. 2013), Ontario (Urbach et al. 2014) and even Michigan (Reames et al. 2015). In Michigan and the UK Health Foundation's Safer Patients Initiative (Benning et al. 2011) it was unclear whether safety improved for longer-term reasons independent of Collaboratives.

This pattern suggests that Collaboratives' effects are highly context-dependent (Vanderkruik and McPherson 2017). However when past studies reported the context at all (Wells et al. 2018) it was often defined broadly

> 'as anything not directly part of the technical QI process that includes the QI methods
> themselves and the clinical interventions. Therefore, context may include factors
> relating to the characteristics of the organizational setting, the individual, his or her role
> in the organization, and the environment' (Kaplan et al. 2010).

Past studies tend to group these factors as:

1. <u>Site suitability</u>: low baseline levels of quality or safety (Franco and Marquez 2011; Algurén

et al. 2019); care groups for whom evidence-based interventions are known but not universally applied (Øvretveit et al. 2002); the safety and/or QI targets are already known to be achievable (Hulscher, Schouten, and Grol 2009) and to which the proposed new work-processes are relevant (Astbury and Leeuw 2010; Shortell et al. 2004).

2. Clinical team composition and size: Clinical team members are judiciously recruited (Singer and Vogus 2013) for having the relevant skills, expertise, and 'connectivity to target populations' (Mohr, Batalden, and Barach 2004, p.37), hence the training and clinical supervision necessary for EBP (Bartholomew et al. 2007; Raghavan et al. 2007); and sufficient in number.

3. Clinical team organisation: Clinical teams have strong relational coordination internally (Gittell, Fairfield, and Bierbaum 2000), but also strong, adaptive team leaders with good communication and task coordination skills (Manser 2009) which enable team learning (Edmondson, Bohmer, and Pisano 2001).

4. Motivation: Clinicians' support (Hulscher, Schouten, and Grol 2009) and team-wide consensus on the QI goals (Bion et al. 2012; Mohr, Batalden, and Barach 2004), with QI being strategically important to their employing organisation (Hulscher, Schouten, and Grol 2009)) and the latter being under pressure to increase productivity (Flin et al. 2000; Pizzi, Goldfarb, and Nash 2001). Motivation is often attributed to a 'positive' (Krein et al. 2010), 'balanced' (Bosch et al. 2008) or 'proficiency' culture (Williams et al. 2017) which provides the psychological safety (Gittell 2000; Garvin, Edmondson, and Gino 2008; Carmeli and Gittell 2009) to acknowledge errors arising from bad luck or unforeseen circumstances.

5. Fit: existing work-processes can accommodate the QI activity, or be adapted to (Pronovost et al. 2006; Bartholomew et al. 2007; Raghavan et al. 2007). Relevant work-processes include workflow, management (e.g. medical record) systems (Conn et al. 2015; Russ et al.

2015), risk management systems (Hudson 2003; Flin et al. 2000), payment systems (Raghavan et al. 2007) and models of care (e.g. the Chronic Care Model) (Hulscher, Schouten, and Grol 2009).

6. Network characteristics: Collaboratives often depend heavily on key individuals who 'transmit information, bridge disparate groups, liaise across parts of networks and enable social and professional interaction' (Cunningham et al. 2011; see also Rossiter et al. 2017) and on collaboration between providers (Raghavan et al. 2007; Kaplan et al. 2010).

7. Material inputs (Bartholomew et al. 2007; Raghavan et al. 2007; Williams et al. 2017) including IT systems which can supply clinical teams with QI data (Dixon-Woods, Baker, et al. 2013; Reames et al. 2015; Vos et al. 2010),

8. Active managerial support (Gioia and Dziadosz 2008), which may depend on who owns the provider organisation (Aarons, Sommerfeld, and Walrath-Greene 2009).

Nevertheless it remains true that

"Key limitations in the existing literature were the lack of a practical conceptual model, the lack of clear definitions of contextual factors, and the lack of well-specified measures' (Kaplan et al. 2010).

More generally, realism still largely lacks a 'context theory' explaining how contexts, at all levels of analysis, may both emerge from and contribute to shaping organisational phenomena and relationships (Bamberger 2008).

The above groups of factors have four types of effect on the mechanisms through which Collaboratives work, which suggests the following taxa of contexts.

1. Motivational:

(a) Actors' prior motivation for achieving the intended outcomes, including safety culture and managerial support.

(b) Outcome 'headroom' i.e. the stipulated outcomes are non-trivial improvements to existing quality and safety.

2. Resource:

(a) Sufficient actors, competent to enact the mechanism(s).

(b) Material inputs, including IT, sufficient for the mechanism to operate.

(c) Finance

3. Collective action:

(a) Structures: networks, hierarchies, markets and organisations through which the actors can coordinate their activity and so jointly implement the policy mechanism.

(b) Policy compatibility. Other, co-existing or prior policy mechanisms and their implementation do not frustrate the focal policy mechanism. The intended outcomes and mechanisms are compatible with the actors' existing activities and priorities.

This paper therefore aims to contribute to the evidence-based development of a taxonomy which realist evaluations can use to differentiate contexts according to the ways in which they impact upon policy mechanisms (and so help disambiguate a key term in evaluations), aid the pooling and comparison of findings across realist evaluations, indeed across policy research generally, and facilitate future realist evaluations. Taking English PSCs as an evidential starting-point we therefore examine:

RQ1. What contextual barriers or facilitators did the implementation of English PSCs encounter?

RQ2. How can these contexts be grouped and classified, in terms of their implications for the mechanisms comprising PSCs?

An initial research study, described below, addressed the first question. Many of its findings concerned the contexts in which PSCs were established and operated, to understand which we developed the second question, on which the present paper focuses. We also consider the implications for conducting realist evaluations.

3   Methods

We addressed these questions by a two-stage method. The first was the Patient Safety Collaborative Evaluation Study (The PiSCES Study), a realist evaluation of the set-up and first three years' work of English PSCs (for short, the 'original evaluation' below). Its full research report is available from the authors. It elicited the prospective programme theory underlying PSCs (Figure 1) and compared it with the actual implementation of PSCs, including the contexts which affected the mechanisms that PSCs used for influencing clinical practice. As noted, policy makers' prospective programme theories may make erroneous assumptions. An empirically elaborated and corrected set of CMOCs for the English PSCs (Figure 2 below) was therefore an output from the original evaluation. It answered RQ1. Specific to the present analysis, the second stage compared these observed CMOCs, in particular their contexts and the consequences of these contexts, with the initial taxonomy outlined above, revising the latter in light of the empirical findings. This answers RQ2.

To produce empirically corrected CMOCs the original evaluation made a purposive, qualitative, maximum-variety sample of study sites and informants so that data collection followed the sequence of PSC implementation as it cascaded downwards through three nested organisational

levels:

1. Regional network i.e. the inter-organisational network that constituted a PSC, comprising its coordinating body, host organisation (the local Academic Health Science Network (AHSN)) and the provider organisations which participated in each PSC. We collected data from all 15 regional PSCs (a census not a sample).

2. Provider organisation In three regions we collected more detailed data about interactions between PSCs and their member-organisations. These regions were purposively sampled as early, middle-term and late adopters of PSC activities with the aim of revealing what contexts promoted or obstructed implemented those activities, and selected on national-level NHS informants' advice about which regions then (2015) had the most and least developed PSCs. In these three we sampled the range of participating providers which included acute hospitals, general practices, and community mental health services (Table 1). This was purposive sampling of a maximum variety of services, improvement infrastructures, geographical conditions and pre-existing networks, aimed at revealing the full range of contexts affecting PSC implementation.

[TABLE 1 HERE]

At provider level we interviewed the senior manager (N=14: some medical, some non-medical) responsible for PSC activity there.

3. Clinical team. Within these providers we sampled clinical team members (N=31) directly providing patient care and participating in PSC activities (which only certain clinical teams did, in each provider). Again this was a purposive sample of informants selected for first-hand knowledge of the mechanisms that PSCs used for changing working practices, and their contexts, so these informants were 'front-line' staff (nurses and doctors predominantly).

During December 2015 – March 2018 we collected data through key informant interviews, a survey of clinical teams, re-analysis of routinely published administrative data, and analysis of 'grey' policy, guidance and managerial documents (Table 2).

[TABLE 2 HERE]

The complete interview schedule is in the research report but, briefly, it was semi-structured by topic. It covered the PSCs' purpose and intended outcomes; origins; implementation; experience and impacts of the PSC; financial management; and, to check completeness, final open questions asking whether the other questions had missed anything important about the PSC. We piloted it by interviewing an AHSN lead, and slightly revised it in consequence. The first two transcripts were independently coded. Three researchers checked the coding for consistency, then agreed explicit coding criteria to remove ambiguities and increase coding reliability. Interviews were by telephone (duration 25-110 minutes), audio recorded, transcribed, cleaned and anonymised. All interviewees had the opportunity to see and correct their transcript. For each site we kept fieldwork notes.

The patient safety climate survey used the Safety, Communication, Operational Reliability and Engagement (SCORE) instrument (Safe & Reliable Healthcare 2014) whose psychometric properties are reported elsewhere (Adair et al. 2018; Profit et al. 2014; Sexton et al. 2006; 2017). It includes an open free-text question eliciting any further comments about the Collaborative. Any NHS site could volunteer for it, so the survey sites were self-selected. During the original evaluation period 72 sites, including general practices, acute and mental health hospital wards, completed a first round of the survey. In practice follow-up surveys are usually made about a year later but since many sites only started surveying in 2017, only six of them had repeated the survey by the end of the evaluation, although by early 2020 many more had done so including sites outwith the original evaluation. So the disparity between first and second round numbers was an artefact of

the original evaluation's timing. Within each clinical team the survey was a census (range 9-151 individuals). Table 2 shows which England-wide, trust-level, routinely- collected outcome data we analysed.

To analyse these data and discover the nature of the CMOCs that constituted PSCs we first triangulated the primary data listed in Table 2. Next we coded them using predetermined codes based on the prospective programme theory assumptions (Figure 1). By comparing our emerging empirical findings across sites, we refined the codes and sub-codes (Bradley, Curry, and Devers 2007), then used the coded data to populate a framework analysis (Gale et al. 2013) also structured according to the prospective programme theory assumptions. For each mechanism (solid arrow in Figure 1) in each CMOC, we systematically compared across sites what actors and activities the PSCs involved, where and when, the intended consequences, whether or how the contexts mentioned in the prospective programme theory had assisted or impeded achieving then, and the actual consequences. For data which did not fit into the existing framework and codes we used first open then axial coding to make an inductive analysis and add categories. We compared the combined findings systematically with the prospective programme theory assumptions, in particular about the contexts and their consequences in each CMOC. None of the evidence contradicted policy-makers' original assumptions about mechanisms or contexts (Figure 1) but it did necessitate elaborating the prospective programme theory to fill in some intermediate outcomes and (still more) additional missing contexts which the prospective programme theory had not foreseen. The original evaluation thus produced a revised, more elaborated and strongly evidence-based retrospective explanation the eight main CMOCs that constituted the English PSCs (Figure 2).

The second stage relied on the first-stage findings, sample and data. We took the first-stage findings about what factors had potentiated or frustrated the mechanism in each CMOC, and how each such context did so. We grouped these explanations by similarity of the causal processes

(generative mechanism) involved, then compared the resulting groups with the initial taxonomy above, revising the latter as necessary. Following good taxonomic practice we tried to formulate taxa that were comprehensive (whilst avoiding the residual category of 'other'), mutually exclusive, unambiguous and theory-based. To meet the latter requirement we categorised the contexts by their effects on the respective mechanism(s). Because we wanted to find taxa that might be qualitatively (analytically, theoretically) generalisable to other settings, we categorised these effects in the most generic causal terms that the data would bear: terms that would accommodate local variations in those contexts but without losing their specific causal character (Eisenhart 2009). These descriptions abstracted from particular organisations, settings and events. For instance when an organisation could not afford activities we categorised that context as 'insufficient budget' rather than (say) 'Dr. X keeping within the 2017 NHS financial limits for hospital Y'.

The NHS Research Ethics Committee system approved the study (reference 15-NI-0235) subject to research governance approval from each study site and informant anonymity.

4   Contexts and their consequences

Our analysis of PSC contexts started from the findings of the original evaluation, where the reader can find fuller details, especially interview statements. The second stage analysis revealed the contexts, and their consequences, for the eight main CMOCs that comprised the PSCs (Figure 2). Solid arrows and boxes in Figure 2 indicate mechanisms, each numbered. Dotted lines and boxes show their respective contexts which we report CMOC by CMOC, noting also how likely each context is, to apply to other kinds of policy mechanism.

[FIGURE 2 HERE]

CMOC 1: Network bodies helps providers to become learning organisations: NHS Improvement (NHSI) established regional PSC coordinating bodies which were to implement three parallel mechanisms (Mechanisms 1,2,3 in Figure 2). One was to help providers become learning organisations in respect of QI and safety. A coordinated network of service providers would identify, train and support patient safety 'champions' in each provider, with the outcome that provider managers actively supported and helped implement PSC activities within their own organisation. Two main contexts affected this mechanism. At national level the Department of Health (DH) and NHSI supported the networks over several years by publicly endorsing the PSCs, monitoring NHS providers' quality and safety activity, practically supporting and financing the networks. That reinforced Mechanism 1 by legitimating PSC activities to providers, indicating to managers that these activities were important and being monitored, so PSCs could not just be ignored. Additionally, NHSI required each AHSN to host the Collaborative for its region. This hosting arrangement gave PSCs a location (setting) from which to operate and a geographically-defined territory, but neither potentiated nor weakened the three main mechanisms that the network coordinating bodies used. PSCs also faced pressures 'from the top' to show within two or three years that they had produced practical results, even cost-savings. Many informants interpreted those pressures as presaging a decision to scrap PSCs if the results did not materialise, although in the event no such decision occurred. These demotivating effects of these uncertainties did however expose the importance avoiding any premature interruption (discontinuity) in the policy support. A further context was that the financial support was limited and fixed-term, which constrained how much training and other activities the network coordinating bodies could undertake. Towards the end of the original evaluation period it made the networks temporarily uncertain about what their future would be. The reinforcing effects of continued government support and financing would appear to apply to almost any policy mechanism.

CMOC 2: Network bodies create inter-organisational networks of clinical teams: Mechanism 2 was that each network would recruit clinical teams, link them across organisations and provide training and expertise to support them so that (as an outcome) the teams could share practical experience, knowledge and data. However the networks varied considerably in how many teams were initially available. In some regions pre-existing networks of clinical teams with a history of collaboration were already undertaking activities like those of Mechanism 2 and already had the relevant knowledge, skills and experience. In these contexts Mechanism 2 therefore could be (and was) established by re-badging something already occurring. Other regions had to establish all this from scratch, making Mechanism 2 slower, harder and more uncertain to establish and operate. Access to networked resources is a context seems likely to apply to at least some other policy mechanisms (Hill 2003), depending on what balance between network, market or hierarchies their implementation depends on. A second context was the financial limitations that restricted the extent of Mechanism 2 as it did for Mechanism 1.

CMOC 3: Network bodies establish cross-organisational measurement systems: The network bodies were also to establish cross-organisational systems for measuring quality, safety and activities to improve them. Endorsement and practical support from the DH and NHS was a context which reinforced Mechanism 3 for the same reasons as Mechanism 1. An important context for Mechanism 2 was how the measurement system was procured. Concurrent marketisation policies stipulated external competitive tendering. Thus one context, concurrent polices, created another: market procurement. That delayed setting up Mechanism 2 for nearly the whole original evaluation period, during which PSCs therefore lacked what policy statements said was an important mechanism for their work. To work around this difficulty at least one PSC meantime devised its own measurement system (Life QI: https://www.lifeqisystem.com/) for sharing project data between clinical teams, and offered it to other PSCs. Once it appeared, the official national system had partly

duplicated this work. It was not of course the first England-wide NHS monitoring system. Our informants repeatedly mentioned earlier NHS performance monitoring systems which increased their data collection work and had been used to blame staff for performance shortfalls. This historical experience of a 'blame culture' made them wonder whether the new measurement systems would have similar consequences. That tended to weaken their motivation actively to help construct or use the QI and safety measurement mechanisms. Insofar as policy mechanisms rely on inputs procured from markets, the procurement context is likely to apply more widely. The after-effects of similar predecessor policies may also (depending on the case) weaken or reinforce later policy mechanisms. Haeder (2012), Knuth (2009) and others corroborate the sometimes path-dependent nature of policy implementation.

CMOC 4: Within providers, learning-organisation activities help strengthen safety climate: Our informants tended to say 'culture' when describing what organisational researchers call 'climate' but either way Mechanism 4 was that senior managerial support within each provider, the recruitment of safety champions, and training staff in quality and safety improvement methods would have the outcome of increasing clinical team members' motivation and capacity to undertake QI and safety-improvement activities. One context affecting this mechanism was managers' judgement of the scope for improving service quality and safety. Managers who thought that they were already doing what the PSCs were proposing had little motivation for further effort in those directions nor, therefore, to allocate staff, money and other resources to support PSC activity at clinical team level:

> *"He* [chief executive] *literally said: 'Everyone in my organisation knows improvement* […]. *This is a waste of money, we don't need to do it'"* (PSCL1).

The number of staff was a context which made the operation of Mechanism 4 across a whole provider require less repetition and coordination, hence less time, in smaller organisations. The

latter also had fewer resources, but network resources (Mechanism 2) compensated. (Again, one context affected another.) Only in general practices did incentives in providers' contracts affect the adoption of QI activities (as policy-makers had assumed), because only in those small organisations were some clinical team members (the GP partners who co-owned the practice) personally exposed to the incentives. The contractual incentives facing larger providers (hospitals, community health services) indirectly affected top managers through performance-related pay, but not clinical team members. During 2015-18 many NHS providers merged, including one of our general practice study sites. There, the consequent re-structuring created the opportunity to reconstitute its QI activities, which increased the effects of Mechanism 4 despite the merged organisation becoming larger (though still far smaller than a hospital). More generally, though, mergers are infrequent and may or may not affect how policy mechanisms work at clinical team level. If not, an organisational setting (organisational identity) changes but not the causal context of those mechanisms.

Because it included the dissemination of new policies, recruiting QI champions and senior clinician supporters, and persuading clinical teams to join the QI programme and training, Mechanism 4 required time to take effect. A negative contextual requirement, not to be taken for granted given the constant organisational 'churn' that other concurrent policies created in the NHS (another example of one context creating another), was therefore continuity: that Mechanism 4 be left undisturbed long enough to work.

Certain concurrent policies conflicted with Mechanism 4. During the evaluation period, budgetary control was an NHS policy priority, closely followed by meeting access (waiting time) targets and ensuring that winter demand pressures did not force EDs or hospital wards to close their doors to new patients. In combination these policies produced staff shortages in many of our study sites, which forced hospital managers and staff to prioritise these other policies over PSC activity:

"*Whenever we go into a particularly tense or difficult period of time on a performance*

basis, so we might slip into these sort of high escalation black alert periods, all of the

safety and quality meetings […] are cancelled" (PSCL6).

A concomitant was the managerial 'performance' or 'blame' climate mentioned above, itself inimical to the safety climate that PSCs required (see Figure 1). Both consequences weakened, even temporarily blocked, Mechanism 4. Against this, though, a number of nurses told us that the enquiries into the hospital failings in Mid-Staffordshire and ensuing policy responses had re-focussed health-workers' minds on QI and patient safety.

A motivational context was the extent to which individual staff (especially practitioners and middle managers) were psychologically receptive to the Collaboratives' ideas about the need and methods for quality and safety improvement. Some saw those ideas as implicitly criticising their own competence:

> "*A lot of people I've found in health-care are incredibly scared of change […] they don't see it that the system could be improved, they see it that they need to be improved*" (PSCOL6).

What many informants described as colleagues' 'resistance to change' determined the extent and speed with which clinical teams adopted these ideas, attitudes and motivations: that is, the safety climate which policy-makers expected Mechanism 4 to instil. Mechanisms which depended on recruiting a certain number or specific mix of participants (e.g. to include doctors) were especially sensitive to this context.

Headroom, or lack of it, for achieving the focal outcomes appears relevant to any policy mechanism (Jung and Lee 2016). Non-healthcare studies (e.g. Zammuto and O'Connor (1992); Jung and Lee (2016)) also analyse the ambivalent ways in which organisational size affects organisational climate. The requirements for continuity, non-conflict with other policies, and actors' motivational support (Kennedy and Fiss 2009) all appear relevant to policy mechanisms generally.

CMOC 5: Changed safety climate changes clinical teams' working practices: As Mechanism 5, clinical teams' strengthened motivation and knowledge of QI work would lead them to use PDSA cycles and team learning to make (as an outcome) their clinical practice more evidence-based. Because its effects accumulated from repeated PDSA cycles, each requiring some months to complete, sufficient continuity without external disruption was also an important context for this mechanism. Second, staff shortages and overload restricted, even halted, Mechanism 5. Individual team members required enough working time to put their QI learning and methods into practice, and backfill for colleagues' QI training. When barely large enough to do their normal work, clinical teams lacked the margin of spare capacity to make innovations. They then saw little practical scope for quality improvement in the present circumstances, despite its feasibility elsewhere, with the same motivational consequences as for Mechanism 4. (Again, one context produced another.) The shortages led to staff overload and so increased staff turnover, which exacerbated the shortages further whilst posts were held vacant (in some cases permanently). With departing staff went QI knowledge, skills, and experience which their replacements had to acquire anew. But the staff shortage context was itself the product of another context: the concurrent policies that also (and more directly) affected Mechanism 4. Outside healthcare, Damanpour (1991) and others have found that innovating requires 'organisational slack' which includes staffing.

CMOC 6: Cross-organisational teamwork helps change working practices: PSC policy assumed that the inter-organisational networks described above draw upon the practical experience and data of other clinical teams' similar QI activities, with the outcome of helping each clinical team change its working practices. Mechanism 6 was thus another mechanism that already-established working relationships within and between the clinical teams helped make more effective. For these concurrent and longer-standing working relationships helped develop trust between the teams

(confidence that admitting quality and safety problems would bring help not blame) and the collaborative invention of feasible changes to working practices. Through the network each clinical team could draw upon any experience which the other teams already had, of devising and implementing technical and organisational remedies for similar quality and safety problems. These conditions helped the clinical teams within providers to undertake and exploit PDSA cycles. A second favourable context was the long-term NHS policy of promoting evidence-based medicine. That gave strong prior legitimacy to what the clinical teams were trying to do, and ready-made technical knowledge for them to apply, which reinforced Mechanism 6, especially in the specialities (e.g. surgery rather than mental health) where EBM was more fully developed. Despite having counterparts in some other sectors, EBM appears the least generalisable of the contexts we found.

CMOC 7: Cross-organisational measurement systems help change clinical working practices: As noted, these systems only appeared belatedly. Mechanism 7 was intended to help clinical teams share data routinely and systematically so that (outcome) teams could compare their changes to clinical working practices, and the effects. A context that affected Mechanism 3 also affected Mechanism 7: that of some clinical team members' memory of measurement systems becoming means for the managerial scrutiny of clinical practice, reinforcing a blame culture. The slow set-up of the measurement systems prevented us observing any further contextual influences on Mechanism 7 during the original evaluation period.

CMOC 8: Changed clinical working practices increase safety and quality: Evidence based medicine itself constituted Mechanism 8. If the underlying evidence were sound, fuller use of evidence-based practice would of itself have the outcome of increasing care quality and safety. Four contexts that affected earlier mechanisms also affected this one:

1. non-interference (continuity) for long enough to take effect (cf. Mechanisms 4,5);

2. headroom for non-trivial quality and safety improvements (cf. Mechanism 4);

3. sufficient staff (cf. Mechanism 5) to implement new working practices. (A context derived in turn from another context i.e. concurrent policies.) The right range of staff was important partly to provide the necessary clinical (e,g. medical) skills and so that the whole team consistently implemented the same working practices.

4. experience of previous similar activities (cf. Mechanisms 2,3,6,7).

Concurrent policies supporting evidence-based medicine (EBM) had stimulated the appearance of prolific guidelines and targets. Some of our informants suspected that managers might also use these targets as another means of monitoring and blaming them for whatever problems were discovered:

> *"One of my colleagues told me that a woman who breaks her hip and goes into hospital and has an operation and then is discharged has over, in some hospitals, 90 guidelines that apply to her care. […] all it's doing, I think, is providing rope for clinicians to […] hang themselves with"* (PSCFL28).

Such comments lead us to speculate (but with some corroboration (Checkland and Harrison 2010)) whether staff therefore applied guidelines and targets selectively, which would constrain the effects of Mechanism 8. Finally, a changed work-process may depending on the case require more or different material inputs (equipment, consumables, beds etc.) than before, and changes to the layout, even size, of hospital and other buildings: a material context which inhibited some attempts to apply evidence-based guidance. After this study finished, hospitals' attempts to respond to CoViD19 provided many further examples. This context would seem to apply to any policy mechanism that depends on material inputs.

5   Discussion and conclusion

Combining and de-duplicating this account of contexts across all eight CMOCs, and then grouping and classifying the contexts in terms of their generic causal implications for PSC mechanisms, produced the categorisations shown in Table 3.

[TABLE 3 HERE]

Some of these taxa are consistent with the kinds of context that (Vanderkruik and McPherson 2017) report and Coldwell's (2019) account of contexts as dynamic and historically located ('Temporal', in Table 3), agentic and immanent (cf. 'Motivational') and relational (cf. Resource contexts). Byng et al. (2008) describe how concurrent policies interact. Kaplan et al.(2012) reported resource availability, capability for improvement ('headroom') and motivation as contexts favouring quality improvement work. Newton et al. (2003), Pettigrew, Ferlie, and McKee (1992) and others analyse absorptive capacity for, or receptivity to, innovation. Singer and Vogus (2013) report size as a context. To our initial taxonomy the findings above add temporal contexts. Indeed in PSCs the temporal contexts (continuity; history; convergent policy) were those which influenced the greatest number of mechanisms (Figure 2, Table 3).

The above categorisations apply irrespective whether each context reinforced, restricted or frustrated operation of the corresponding mechanism(s), hence irrespective of what outcomes the focal mechanism actually produced. To keep the taxa mutually exclusive, policy support is defined as excluding financial support (included under 'Financial' in Table 3). 'Network' refers to contemporary not past mechanisms ('History'); one difference is that the effects of contemporary institutions may be expected to continue, those of former mechanisms to weaken, over time. 'Market' refers to external contractual relationships, whether procurement or contractual income

and incentives, and includes quasi-markets. 'Hierarchy' includes quasi-hierarchies, the vertical inter-organisational links through which governments exercise governance over arms-length bodies (e.g. NHS organisations) (Exworthy, Powell, and Mohan 1999). Convergent contexts may include an unconnected policy mechanism which usurps the one that policy-makers initiated and makes the policy succeed for unforeseen reasons (Marshall et al. 2000). Some contexts were concatenated. Other policies concurrent with the policy of establishing PSCs restricted hospital spending. That caused staff and other input shortages, and in some hospitals reduced the outcome headroom for PSCs. Other concurrent policies stipulated market procurement (for PSCs, of measurement systems). Complex though it already is, Figure 2 omits these further concatenations. Each context in these chains might affect multiple mechanisms (see Table 3). Multiple contexts might affect each mechanism (Figure 2). In PSCs, the size of a provider organisation (an organisational context) determined how far contractual incentives (a market context) affected the clinical teams. One context might thus increase or decrease the effect of another, becoming, so to speak, the context-of-a-context. The fact that the implementers recognised a context (e.g. heavy demand for hospital care) as important did not necessarily mean that they could influence it.

As a census of PSCs, our original original evaluation had neither selection nor participation bias at national and regional levels. Its remit compelled us to select sites where PSC activities were already proceeding, omitting other sites and therefore contexts, so we may under-report the range of contexts unfavourable to launching PSCs. Fieldwork covered PSCs' first three years, not any subsequent contextual changes.

Our conclusions suggest some key points for evaluators to consider when conducting evaluations. One is to understand contexts as generative, causal structures rather than as social locations, places or times. Similar factors (e.g. material resources) may function as outcomes for some policies (e.g. privatisation), mechanism for others (e.g. automation) and as context for others

again (e.g. Collaboratives), as emerges when comparing studies (Williams et al. 2017; Jody Hoffer Gittell et al. 2008; Kaplan et al. 2010; Coldwell 2019). Contexts may be social, physical or both (e.g. epidemiological (Dixon-Woods 2014)) depending on the case, but the term 'context' properly applies only to factors which, an evaluation finds, causally affect the focal intervention's capacity to produce the outcomes which policy-makers intended (irrespective of other outcomes or side-effects). Other local circumstances are settings not contexts. One way to particularise or localise an evaluation is therefore to describe not its setting but what specific combination of contexts the evaluators found (cf. Simons 2015). Differences in setting therefore do not in principle prevent wider use of the above taxonomy. The latter would appear to have internal validity in the sense of applicability to similar mechanisms elsewhere (Eisenhart 2009), because other patient safety collaboratives use essentially similar mechanisms to the PSCs and are therefore susceptible to the same disrupting or potentiating factors. Indeed our taxa derive partly from findings about other patient safety collaboratives.

As explained, many of the above contexts (except EBM) may be more widely relevant. Nevertheless one might expect external generalisability of these taxa to decrease, the less similar to PSCs other policy mechanisms are. The above taxa might apply with only slight adjustments to other healthcare networks (e.g. professional networks, care networks (Southon, Perkins, and Galler 2005)) but require larger additions or revisions for application to other sectors (e.g. networks of industrial cooperatives) or other governance structures (e.g. policy implementation through markets not networks). Table 3 provides a list of contexts to check for in evaluations, but is probably not exhaustive. Although contexts are real, a taxonomy of them is only a theory-based way of analysing evaluation data and using them for policy development. To categorise a context as an instance of one of the above taxa is to predict the ways in which it affects the mechanism(s) that it bears upon. Testing such claims, and revising the corresponding taxonomy as necessary, is something that

evaluations seldom do although many of them readily could Eisenhart (2009), provided that their sampling strategy included sites with similar mechanisms but apparently contrasting contexts.

Comparing Figures1 and 2 shows how Patient Safety Collaboratives in England depended on more contexts, and more complex contexts, than policy-makers first recognised. Such complexities make the outcomes of a given policy mechanism highly contingent (Sayer 1992) and make policy highly fallible (Kaneko 1999). They make it hard to formulate valid, sufficiently comprehensive assumptions about how and under what conditions a policy will produce its intended outcomes: a valid 'programme theory' (Pawson 2008), 'theory of change' (Mason and Barnes 2007) or 'change model' (Chen 2014). This study contributes to alleviating that difficulty by newly identifying time-dependent contexts and systematically listing other types of context that future Collaboratives, and perhaps other policy initiatives, might adjust for (Dixon-Woods 2014).

Acknowledgements

References

Aarons, Gregory, David Sommerfeld, and Christine Walrath-Greene. 2009. 'Evidence-Based

Practice Implementation: The Impact of Public versus Private Sector Organization Type on

Organizational Support, Provider Attitudes, and Adoption of Evidence-Based Practice'.

Implementation Science 4 (1): 83. https://doi.org/10.1186/1748-5908-4-83.

Adair, Kathryn C., Krystina Quow, Allan Frankel, Paul J. Mosca, Jochen Profit, Allison Hadley,

Michael Leonard, and J. Bryan Sexton. 2018. 'The Improvement Readiness Scale of the

SCORE Survey: A Metric to Assess Capacity for Quality Improvement in Healthcare'. *BMC

Health Services Research* 18 (1): 975. https://doi.org/10.1186/s12913-018-3743-0.

Algurén, Beatrix, Annika Nordin, Boel Andersson-Gäre, and Anette Peterson. 2019. 'In-Depth

Comparison of Two Quality Improvement Collaboratives from Different Healthcare Areas

Based on Registry Data—Possible Factors Contributing to Sustained Improvement in

Outcomes beyond the Project Time'. *Implementation Science* 14 (1): 74.

https://doi.org/10.1186/s13012-019-0926-y.

Anon. 2013. 'Amélioration De La Qualité Et De La Sécurité Des Soins'. Haute Autorité de Santé.

Astbury, Brad, and Frans L. Leeuw. 2010. 'Unpacking Black Boxes: Mechanisms and Theory

Building in Evaluation'. *American Journal of Evaluation* 31 (3): 363–81.

Aveling, Emma-Louise, Michael Parker, and Mary Dixon-Woods. 2016. 'What Is the Role of

Individual Accountability in Patient Safety? A Multi-Site Ethnographic Study'. *Sociology of

Health & Illness* 38 (2): 216–32. https://doi.org/10.1111/1467-9566.12370.

Bamberger, Peter. 2008. 'Beyond Contextualization: Using Context Theories to Narrow the Micro-

Macro Gap in Management Research'. *Academy of Management Journal* 51 (5): 839–46.

Barceló, Alberto, Elizabeth Cafiero, Melanie de Boer, Alejandro Escobar Mesa, Marcelina García

Lopez, Rosa Aurora Jiménez, Agustín Lara Esqueda, José Antonio Martinez, Esperanza

Medina Holguin, and Micheline Meiners. 2010. 'Using Collaborative Learning to Improve

Diabetes Care and Outcomes: The VIDA Project'. *Primary Care Diabetes* 4 (3): 145–53.

Barnes, Marian, Elizabeth Matka, and Helen Sullivan. 2003. 'Evidence, Understanding and

Complexity: Evaluation in Non-Linear Systems'. *Evaluation* 9 (3): 265–84.

Bartholomew, N. G., G. W. Joe, G. A. Rowan-Szal, and D. D. Simpson. 2007. 'Counselor Assessments of Training and Adoption Barriers.' *Journal of Substance Abuse Treatment* 33 (2): 193–99. https://doi.org/10.1016/j.jsat.2007.01.005.

Bastien, Joëlle. 2009. 'Goal Ambiguity and Informal Discretion in the Implementation of Public Policies: The Case of Spanish Immigration Policy'. *International Review of Administrative Sciences* 75 (4): 665–85.

Benn, Jonathan, Susan Burnett, Anam Parand, Anna Pinto, and Charles Vincent. 2012. 'Factors Predicting Change in Hospital Safety Climate and Capability in a Multi-Site Patient Safety Collaborative: A Longitudinal Survey Study'. *BMJ Quality & Safety*, May, bmjqs-2011-000286. https://doi.org/10.1136/bmjqs-2011-000286.

Benning, Amirta, Maisoon Ghaleb, Anu Suokas, Mary Dixon-Woods, Jeremy Dawson, Nick Barber, Bryony Dean Franklin, et al. 2011. 'Large Scale Organisational Intervention to Improve Patient Safety in Four UK Hospitals: Mixed Method Evaluation'. *BMJ (Clinical Research Ed.)* 342 (February): d195.

Bergen, A., and A. White. 2005. '"Implementation Deficit" and "Street-Level Bureaucracy": Policy, Practice and Change in the Development of Community Nursing Issues'. *Health and Social Care in the Community* 13 (1): 1–10.

Berwick, D. 2013. 'A Promise to Learn – a Commitment to Act. Improving the Safety of Patients in England'. London: National Advisory Group on the Safety of Patients in England.

Bion, Julian, Annette Richardson, Peter Hibbert, Jeanette Beer, Tracy Abrusci, Martin McCutcheon, Jane Cassidy, Jane Eddleston, Kevin Gunning, and Geoff Bellingan. 2012. '"Matching Michigan": A 2-Year Stepped Interventional Programme to Minimise Central Venous Catheter-Blood Stream Infections in Intensive Care Units in England'. *BMJ Qual Saf*, bmjqs-2012-001325.

Bosch, Marije, Rob Dijkstra, Michel Wensing, Trudy van der Weijden, and Richard Grol. 2008.

'Organizational Culture, Team Climate and Diabetes Care in Small Office-Based Practices'.

BMC Health Services Research 8 (1): 180. https://doi.org/10.1186/1472-6963-8-180.

Bradford-Hill, A. 1965. 'The Environment and Disease: Association or Causation?' *Proceedings of

the Royal Society of Medicine* 58 (5): 295–300.

Bradley, Elizabeth H, Leslie A Curry, and Kelly J Devers. 2007. 'Qualitative Data Analysis for

Health Services Research: Developing Taxonomy, Themes, and Theory'. *Health Services

Research* 42 (4): 1758–72. https://doi.org/10.1111/j.1475-6773.2006.00684.x.

Brewer, Barbara B. 2006. 'Relationships among Teams, Culture, Safety, and Cost Outcomes'.

Western Journal of Nursing Research 28 (6): 641–53.

Broer, Tineke, Anna P. Nieboer, and Roland Bal. 2012. 'Governing Mental Health Care: How

Power Is Exerted in and through a Quality Improvement Collaborative'. *Public

Administration* 90 (3): 800–815. https://doi.org/10.1111/j.1467-9299.2011.01967.x.

Byng, Richard, Ian Norman, Sally Redfern, and Roger Jones. 2008. 'Exposing the Key Functions of

a Complex Intervention for Shared Care in Mental Health: Case Study of a Process

Evaluation'. *BMC Health Services Research* 8 (1): 274. https://doi.org/10.1186/1472-6963-

8-274.

Carmeli, Abraham, and Jody Hoffer Gittell. 2009. 'High-Quality Relationships, Psychological

Safety, and Learning from Failures in Work Organizations'. *Journal of Organizational

Behavior* 30 (6): 709–29. https://doi.org/10.1002/job.565.

Checkland, K, and S Harrison. 2010. 'The Impact of the Quality and Outcomes Framework on

Practice Organisation and Service Delivery: Summary of Evidence from Two Qualitative

Studies'. *Quality in Primary Care* 18 (2): 139–46.

Chen, Huey T. 2014. *Practical Program Evaluation: Theory-Driven Evaluation and the Integrated

Evaluation Perspective*. Sage Publications.

Coldwell, Mike. 2019. 'Reconsidering Context: Six Underlying Features of Context to Improve Learning from Evaluation'. *Evaluation* 25 (1): 99–117.

Conn, Lesley Gotlib, Marg McKenzie, Emily A. Pearsall, and Robin S. McLeod. 2015. 'Successful Implementation of an Enhanced Recovery after Surgery Programme for Elective Colorectal Surgery: A Process Evaluation of Champions' Experiences'. *Implementation Science* 10 (1): 99.

Cunningham, Frances C., Geetha Ranmuthugala, Jennifer Plumb, Andrew Georgiou, Johanna I. Westbrook, and Jeffrey Braithwaite. 2011. 'Health Professional Networks as a Vector for Improving Healthcare Quality and Safety: A Systematic Review'. *BMJ Quality & Safety*, November. https://doi.org/10.1136/bmjqs-2011-000187.

Curry, Leslie A., Marie A. Brault, Erika L. Linnander, Zahirah McNatt, Amanda L. Brewster, Emily Cherlin, Signe Peterson Flieger, Henry H. Ting, and Elizabeth H. Bradley. 2017. 'Influencing Organisational Culture to Improve Hospital Performance in Care of Patients with Acute Myocardial Infarction: A Mixed-Methods Intervention Study'. *BMJ Qual Saf*, bmjqs-2017-006989.

Damanpour, Fariborz. 1991. 'Organizational Innovation: A Meta-Analysis of Effects of Determinants and Moderators'. *Academy of Management Journal* 34 (3): 555–90. https://doi.org/10.2307/256406.

Damschroder, Laura J., and Julie C. Lowery. 2013. 'Evaluation of a Large-Scale Weight Management Program Using the Consolidated Framework for Implementation Research (CFIR)'. *Implementation Science* 8 (1): 51.

Davies, Elizabeth, Dale Shaller, Susan Edgman-Levitan, Dana G Safran, Gary Oftedahl, John Sakowski, and Paul D Cleary. 2008. 'Evaluating the Use of a Modified CAHPS® Survey to Support Improvements in Patient-Centred Care: Lessons from a Quality Improvement Collaborative'. *Health Expectations* 11 (2): 160–76. https://doi.org/10.1111/j.1369-

7625.2007.00483.x.

Department of Health. 2014. 'Hard Truths: The Journey to Putting Patients First: Volume One of the Government Response to the Mid Staffordshire NHS Foundation Trust Public Inquiry'. Cm 8777-1. London: Department of Health. https://socialwelfare.bl.uk/subject-areas/services-activity/health-services/tso/hard13.aspx.

Dillon, Robin L., Catherine H. Tinsley, Peter M. Madsen, and Edward W. Rogers. 2016. 'Organizational Correctives for Improving Recognition of Near-Miss Events'. *Journal of Management* 42 (3): 671–97. https://doi.org/10.1177/0149206313498905.

Dixon-Woods, Mary. 2014. 'The Problem of Context in Quality Improvement'. *Perspectives on Context. London: Health Foundation*, 87–101.

Dixon-Woods, Mary, Richard Baker, Kathryn Charles, Jeremy Dawson, Gabi Jerzembek, Graham Martin, Imelda McCarthy, Lorna McKee, Joel Minion, and Piotr Ozieranski. 2013. 'Culture and Behaviour in the English National Health Service: Overview of Lessons from a Large Multimethod Study'. *BMJ Quality & Safety*, bmjqs-2013-001947.

Dixon-Woods, Mary, Myles Leslie, Carolyn Tarrant, and Julian Bion. 2013. 'Explaining Matching Michigan: An Ethnographic Study of a Patient Safety Program'. *Implementation Science* 8: 70. https://doi.org/10.1186/1748-5908-8-70.

Dückers, Michel LA, Peter Spreeuwenberg, Cordula Wagner, and Peter P. Groenewegen. 2009. 'Exploring the Black Box of Quality Improvement Collaboratives: Modelling Relations between Conditions, Applied Changes and Outcomes'. *Implementation Science* 4 (1): 74.

Edmondson, Amy C., Richard M. Bohmer, and Gary P. Pisano. 2001. 'Disrupted Routines: Team Learning and New Technology Implementation in Hospitals'. *Administrative Science Quarterly* 46 (4): 685–716.

Eisenhart, Margaret. 2009. 'Generalization from Qualitative Inquiry'. *Generalizing from Educational Research: Beyond Qualitative and Quantitative Polarization*, 51–66.

Exworthy, M., M. Powell, and J. Mohan. 1999. 'The NHS: Quasi-Market, Quasi-Hierarchy and Quasi-Network?' *Public Money and Management* 19 (4): 15–22.

Ferlie, E.B., and S.M. Shortell. 2001. 'Improving the Quality of Health Care in the United Kingdom and the United States: A Framework for Change'. *Milbank Quarterly.* 79 (2): 281–315.

Flin, Rhona, Kathryn Mearns, Paul O'Connor, and Robin Bryden. 2000. 'Measuring Safety Climate: Identifying the Common Features'. *Safety Science* 34 (1–3): 177–92.

Franco, Lynne Miller, and Lani Marquez. 2011. 'Effectiveness of Collaborative Improvement: Evidence from 27 Applications in 12 Less-Developed and Middle-Income Countries'. *BMJ Quality & Safety*, bmjqs. 2010.044388.

Fulton, Brent D., Susan L. Ivey, Hector P. Rodriguez, and Stephen M. Shortell. 2017. 'Countywide Physician Organization Learning Collaborative and Changes in Hospitalization Rates'.

Gale, Nicola K., Gemma Heath, Elaine Cameron, Sabina Rashid, and Sabi Redwood. 2013. 'Using the Framework Method for the Analysis of Qualitative Data in Multi-Disciplinary Health Research'. *BMC Medical Research Methodology* 13 (1): 117. https://doi.org/10.1186/1471-2288-13-117.

Garvin, David A., Amy C. Edmondson, and Francesca Gino. 2008. 'Is Yours a Learning Organization?' *Harvard Business Review* 86 (3): 109–16.

Gioia, Deborah, and Gregory Dziadosz. 2008. 'Adoption of Evidence-Based Practices in Community Mental Health: A Mixed-Method Study of Practitioner Experience'. *Community Mental Health Journal* 44 (5): 347–57. https://doi.org/10.1007/s10597-008-9136-9.

Gittell, J.H., K. Fairfield, and B. Bierbaum. 2000. 'Impact of Relational Coordination on the Quality of Care, Post-Operative Pain and Functioning, and Length of Stay: A Nine-Hospital Study of Surgical Patients'. *Medical Care* 38 (8): 807–19.

Gittell, Jody Hoffer. 2000. 'Organizing Work to Support Relational Co-Ordination.' *International Journal of Human Resource Management* 11 (3): 517–39.

https://doi.org/10.1080/095851900339747.

Gittell, Jody Hoffer, Dana Weinberg, Susan Pfefferle, and Christine Bishop. 2008. 'Impact of Relational Coordination on Job Satisfaction and Quality Outcomes: A Study of Nursing Homes'. *Human Resource Management Journal* 18 (2): 154–70.

Haeder, Simon F. 2012. 'Beyond Path Dependence: Explaining Healthcare Reform and Its Consequences'. *Policy Studies Journal* 40: 65–86.

Hatahet, M. Ammar, J. Bowhan, and Elizabeth A. Clough. 2004. 'Wisconsin Collaborative for Healthcare Quality (WCHQ): Lessons Learned'. *WMJ-MADISON-* 103: 45–48.

Herepath, Andrea, Martin Kitchener, and Justin Waring. 2015. 'A Realist Analysis of Hospital Patient Safety in Wales: Applied Learning for Alternative Contexts from a Multisite Case Study'. *Health Services and Delivery Research* 3 (40).

Hill, Heather C. 2003. 'Understanding Implementation: Street-Level Bureaucrats' Resources for Reform'. *Journal of Public Administration Research and Theory* 13 (3): 265–82. https://doi.org/10.1093/jpart/mug024.

Hudson, Patrick. 2003. 'Applying the Lessons of High Risk Industries to Health Care'. *BMJ Quality & Safety* 12 (suppl 1): i7–12.

Hulscher, Marlies, Loes M. T. Schouten, and R. Grol. 2009. 'Collaboratives'. London: Health Foundation.

Jones, Marshall B. 2007. 'The Multiple Sources of Mission Drift'. *Nonprofit and Voluntary Sector Quarterly* 36 (2): 299–307. https://doi.org/10.1177/0899764007300385.

Jung, Chan Su, and Geon Lee. 2016. 'Organizational Climate, Leadership, Organization Size, and Aspiration for Innovation in Government Agencies'. *Public Performance & Management Review* 39 (4): 757–82. https://doi.org/10.1080/15309576.2015.1137764.

Kaneko, Masahiko. 1999. 'A Methodological Inquiry into the Evaluation of Smoking Cessation Programmes'. *Health Education Research* 14 (3): 433–41.

Kaplan, Heather C., Patrick W. Brady, Michele C. Dritz, David K. Hooper, W. Linam, Craig M. Froehle, and Peter Margolis. 2010. 'The Influence of Context on Quality Improvement Success in Health Care: A Systematic Review of the Literature'. *The Milbank Quarterly* 88 (4): 500–559.

Kaplan, Heather C., Lloyd P. Provost, Craig M. Froehle, and Peter A. Margolis. 2012. 'The Model for Understanding Success in Quality (MUSIQ): Building a Theory of Context in Healthcare Quality Improvement'. *BMJ Quality & Safety* 21 (1): 13–20.

Kennedy, Mark Thomas, and Peer Christian Fiss. 2009. 'Institutionalization, Framing, and Diffusion: The Logic of TQM Adoption and Implementation Decisions among US Hospitals'. *Academy of Management Journal* 52 (5): 897–918.

Keogh, B. 2013. 'Review into the Quality of Care and Treatment Provided by 14 Hospital Trusts in England: Overview Report'. London: NHS England.

Kilo, Charles M. 1998. 'A Framework for Collaborative Improvement: Lessons from the Institute for Healthcare Improvement's Breakthrough Series.' *Quality Management in Healthcare* 6 (4): 1–14.

Knight, Andrew W., Dale Ford, Ralph Audehm, Stephen Colagiuri, and James Best. 2012. 'The Australian Primary Care Collaboratives Program: Improving Diabetes Care'. *BMJ Qual Saf*, bmjqs-2011-000460.

Knuth, Matthias. 2009. 'Path Shifting and Path Dependence: Labor Market Policy Reforms under German Federalism'. *Intl Journal of Public Administration* 32 (12): 1048–69.

Krein, Sarah L., Laura J. Damschroder, Christine P. Kowalski, Jane Forman, Timothy P. Hofer, and Sanjay Saint. 2010. 'The Influence of Organizational Context on Quality Improvement and Patient Safety Efforts in Infection Prevention: A Multi-Center Qualitative Study'. *Social Science & Medicine* 71 (9): 1692–1701.

Kringos, Dionne S., Rosa Sunol, Cordula Wagner, Russell Mannion, Philippe Michel, Niek S.

Klazinga, Oliver Groene, and $author firstName $author.lastName. 2015. 'The Influence of
    Context on the Effectiveness of Hospital Quality Improvement Strategies: A Review of
    Systematic Reviews'. *BMC Health Services Research* 15 (1): 277.
    https://doi.org/10.1186/s12913-015-0906-0.

Lewis, Cara C., Meredith R. Boyd, Callie Walsh-Bailey, Aaron R. Lyon, Rinad Beidas, Brian
    Mittman, Gregory A. Aarons, Bryan J. Weiner, and David A. Chambers. 2020. 'A
    Systematic Review of Empirical Studies Examining Mechanisms of Implementation in
    Health'. *Implementation Science* 15 (1): 21. https://doi.org/10.1186/s13012-020-00983-3.

Manley, K., C. Jackson, C. McKenzie, A. Martin, and T. Wright. 2017. 'Safety Culture, Quality
    Improvement, Realist Evaluation (SCQIRE). Evaluating the Impact of the Patient Safety
    Collaborative Initiative Developed by Kent Surrey and Sussex Academic Health Science
    Network (Kssahsn) on Safety Culture, Leadership, and Quality Improvement Capability'.
    Canterbury: Kent, Surrey and Sussex Academic Health Science Network.

Manser, T. 2009. 'Teamwork and Patient Safety in Dynamic Domains of Healthcare: A Review of
    the Literature'. *Acta Anaesthesiologica Scandinavica* 53 (2): 143–51.

Marchal, Bruno, Sara van Belle, Josefien van Olmen, Tom Hoerée, and Guy Kegels. 2012. 'Is
    Realist Evaluation Keeping Its Promise? A Review of Published Empirical Studies in the
    Field of Health Systems Research'. *Evaluation* 18 (2): 192–212.
    https://doi.org/10.1177/1356389012442444.

Mardon, Russell E., Kabir Khanna, Joann Sorra, Naomi Dyer, and Theresa Famolaro. 2010.
    'Exploring Relationships between Hospital Patient Safety Culture and Adverse Events'.
    Journal of Patient Safety 6 (4): 226–32.

Marshall, M.N., P. G. Shekelle, S. Leatherman, and R. H. Brook. 2000. 'Public Release of
    Performance Data. What Do We Expect to Gain?: A Review of the Evidence.' *Journal of the
    America Medical Association* 284:00:00 (14:00:00): 1865–74.

Mason, Paul, and Marian Barnes. 2007. 'Constructing Theories of Change: Methods and Sources'. *Evaluation* 13 (2): 151–70.

Mohr, J., P. Batalden, and Paul Barach. 2004. 'Integrating Patient Safety into the Clinical Microsystem'. *Quality and Safety in Health Care* 13 (suppl 2): ii34–38.

Nadeem, Erum, S. Serene Olin, Laura Campbell Hill, Kimberly Eaton Hoagwood, and Sarah McCue Horwitz. 2013. 'Understanding the Components of Quality Improvement Collaboratives: A Systematic Literature Review'. *The Milbank Quarterly* 91 (2): 354–94.

Nembhard, Ingrid M. 2012. 'All Teach, All Learn, All Improve?: The Role of Interorganizational Learning in Quality Improvement Collaboratives'. *Health Care Management Review* 37 (2): 154.

Newton, John, Joanne Graham, Kevin McLoughlin, and Alice Moore. 2003. 'Receptivity to Change in a General Medical Practice'. *British Journal of Management* 14 (2): 143–53.

Øvretveit, J., Paul Cleary, Shan Cretin, D. Gustafson, K. McInnes, H. McLeod, T. Molfenter, P. Plsek, and Glenn Robert. 2002. 'Quality Collaboratives: Lessons from Research'. *Quality and Safety in Health Care* 11 (4): 345–51.

Øvretveit, John C., Paul G. Shekelle, Sydney M. Dy, Kathryn M. McDonald, Susanne Hempel, Peter Pronovost, Lisa Rubenstein, Stephanie L. Taylor, Robbie Foy, and Robert M. Wachter. 2011. 'How Does Context Affect Interventions to Improve Patient Safety? An Assessment of Evidence from Studies of Five Patient Safety Practices and Proposals for Research'. *Quality and Safety in Health Care* 20 (7): 604–10.

Parand, Anam, Jonathan Benn, Susan Burnett, Anna Pinto, and Charles Vincent. 2012. 'Strategies for Sustaining a Quality Improvement Collaborative and Its Patient Safety Gains'. *International Journal for Quality in Health Care*, mzs030.

Pawson, R. 2008. *Evidence-Based Policy*. London: Sage.

Pawson, R., T. Greenhalgh, C. Harvey, and K. Walshe. 2005. 'Realist Review – a New Method of

Systematic Review Designed for Complex Policy Interventions'. *Journal of Health Services Research and Policy* 10 (Suppl. 1): 21–34.

Pawson, Ray, and Nick Tilley. 1997. *Realistic Evaluation*. London: Sage.

Pettigrew, Andrew, Ewan Ferlie, and Lorna McKee. 1992. 'Shaping Strategic Change-The Case of the NHS in the 1980s'. *Public Money & Management* 12 (3): 27–31.

Pizzi, L., N. Goldfarb, and D. Nash. 2001. 'Promoting a Culture of Safety'. In *Shojania K., Duncan B., McDonald K., & Wachter R. (Eds.) Making Health Care Safer: A Critical Analysis of Patient Safety Practices*, 43:447–52. Evidence Report/Technology Assessment. Rockville: Agency for Healthcare Research and Quality. http://www.ahrq.gov/clinic/ptsafety/chap40.htm.

Profit, Jochen, Paul J. Sharek, Amber B. Amspoker, Mark A. Kowalkowski, Courtney C. Nisbet, Eric J. Thomas, Whitney A. Chadwick, and J. Bryan Sexton. 2014. 'Burnout in the NICU Setting and Its Relation to Safety Culture'. *BMJ Qual Saf* 23 (10): 806–13. https://doi.org/10.1136/bmjqs-2014-002831.

Pronovost, Peter, Dale Needham, Sean Berenholtz, David Sinopoli, Haitao Chu, Sara Cosgrove, Bryan Sexton, et al. 2006. 'An Intervention to Decrease Catheter-Related Bloodstream Infections in the ICU'. *The New England Journal of Medicine* 355 (26): 2725–32. https://doi.org/10.1056/NEJMoa061115.

Raghavan, Ramesh, Moira Inkelas, Todd Franke, and Neal Halfon. 2007. 'Administrative Barriers to the Adoption of High-Quality Mental Health Services for Children in Foster Care: A National Study'. *Administration and Policy in Mental Health and Mental Health Services Research* 34 (3): 191–201. https://doi.org/10.1007/s10488-006-0095-6.

Rasmussen, Jens, and O. M. Pedersen. 1984. 'Human Factors in Probabilistic Risk Analysis and Risk Management'. *Operational Safety of Nuclear Power Plants* 1: 181–94.

Reames, Bradley N., Robert W. Krell, Darrell A. Campbell, and Justin B. Dimick. 2015. 'A

Checklist-Based Intervention to Improve Surgical Outcomes in Michigan: Evaluation of the Keystone Surgery Program'. *JAMA Surgery* 150 (3): 208–15.

Rog, D. 2012. 'When Background Becomes Foreground: Toward Context-sensitive Evaluation Practice'. *New Directions for Evaluation* 2012 (135): 25–40.

Rossiter, Meghan, Jennifer Verma, Jean-Louis Denis, Stephen Samis, Richard Wedge, and Chris Power. 2017. 'Governing Collaborative Healthcare Improvement: Lessons From an Atlantic Canadian Case'. *International Journal of Health Policy and Management* 6 (12): 691–94. https://doi.org/10.15171/ijhpm.2017.60.

Russ, Stephanie J., Nick Sevdalis, Krishna Moorthy, Erik K. Mayer, Shantanu Rout, Jochem Caris, Jenny Mansell, Rachel Davies, Charles Vincent, and Ara Darzi. 2015. 'A Qualitative Evaluation of the Barriers and Facilitators toward Implementation of the WHO Surgical Safety Checklist across Hospitals in England: Lessons from the "Surgical Checklist Implementation Project"'. *Annals of Surgery* 261 (1): 81–91.

Safe & Reliable Healthcare. 2014. 'The Integrated SCORE (Safety, Communication, Operational Reliability & Engagement Survey) Survey for Organizational Learning and Improvement'. Safe & Reliable Healthcare.

Sayer, Andrew. 1992. *Method in Social Science: A Realist Approach*. Psychology Press.

Schouten, Loes M. T., Marlies E. J. L. Hulscher, Jannes J. E. van Everdingen, Robbert Huijsman, and Richard P. T. M. Grol. 2008. 'Evidence for the Impact of Quality Improvement Collaboratives: Systematic Review'. *BMJ* 336 (7659): 1491–94. https://doi.org/10.1136/bmj.39570.749884.BE.

Sexton, J., Robert L. Helmreich, Torsten B. Neilands, Kathy Rowan, Keryn Vella, James Boyden, Peter R. Roberts, and Eric J. Thomas. 2006. 'The Safety Attitudes Questionnaire: Psychometric Properties, Benchmarking Data, and Emerging Research'. *BMC Health Services Research* 6 (1): 44. https://doi.org/10.1186/1472-6963-6-44.

Sexton, J., Stephanie P. Schwartz, Whitney A. Chadwick, Kyle J. Rehder, Jonathan Bae, Joanna Bokovoy, Keith Doram, Wayne Sotile, Kathryn C. Adair, and Jochen Profit. 2017. 'The Associations between Work–Life Balance Behaviours, Teamwork Climate and Safety Climate: Cross-Sectional Survey Introducing the Work–Life Climate Scale, Psychometric Properties, Benchmarking Data and Future Directions'. *BMJ Qual Saf* 26 (8): 632–40. https://doi.org/10.1136/bmjqs-2016-006032.

Shaw, Eric K., Sabrina M. Chase, Jenna Howard, Paul A. Nutting, and Benjamin F. Crabtree. 2012. 'More Black Box to Explore: How Quality Improvement Collaboratives Shape Practice Change'. *The Journal of the American Board of Family Medicine* 25 (2): 149–57.

Sheaff, R. 1997. 'Healthcare Access and Mobility between the UK and Other European Union States: An Implementation Surplus'. *Health Policy* 42 (3): 239–53.

Shortell, Stephen M., Jill A. Marsteller, Michael Lin, Marjorie L. Pearson, Shin-Yi Wu, Peter Mendel, Shan Cretin, and Mayde Rosen. 2004. 'The Role of Perceived Team Effectiveness in Improving Chronic Illness Care'. *Medical Care* 42 (11): 1040–48.

Simons, Helen. 2015. 'Interpret in Context: Generalizing from the Single Case in Evaluation'. *Evaluation* 21 (2): 173–88. https://doi.org/10.1177/1356389015577512.

Singer, Sara J., and Timothy J. Vogus. 2013. 'Reducing Hospital Errors: Interventions That Build Safety Culture'. *Annual Review of Public Health* 34: 373–96.

Smith, Shawna N., M. Todd Greene, Lona Mody, Jane Banaszak-Holl, Laura D. Petersen, and Jennifer Meddings. 2017. 'Evaluation of the Association between Nursing Home Survey on Patient Safety Culture (NHSOPS) Measures and Catheter-Associated Urinary Tract Infections: Results of a National Collaborative'. *BMJ Qual Saf*, bmjqs-2017-006610.

Solberg, Leif I. 2005. 'If You've Seen One Quality Improvement Collaborative …'. *The Annals of Family Medicine* 3 (3): 198–99. https://doi.org/10.1370/afm.304.

Southon, G., R. Perkins, and D. Galler. 2005. 'Networks: A Key to the Future of Health Services'.

Australian Health Review. 29 (3): 317–26.

The Mid Staffordshire NHS Foundation Trust Inquiry. 2010. 'Independent Enquiry into the Care Provided by Mid Staffordshire NHS Foundation Trust 2005-2009'. London: HMSO.

The RAMESES II Project. 2017a. 'What Is a Mechanism? What Is a Programme Mechanism?' NIHR. http://ramesesproject.org/media/RAMESES_II_What_is_a_mechanism.pdf.

———. 2017b. 'What Realists Mean by Context; or, Why Nothing Works Everywhere or for Everyone'. NIHR.

Urbach, David R., Anand Govindarajan, Refik Saskin, Andrew S. Wilton, and Nancy N. Baxter. 2014. 'Introduction of Surgical Safety Checklists in Ontario, Canada'. *New England Journal of Medicine* 370 (11): 1029–38.

Vanderkruik, Rachel, and Marianne E. McPherson. 2017. 'A Contextual Factors Framework to Inform Implementation and Evaluation of Public Health Initiatives'. *American Journal of Evaluation* 38 (3): 348–59. https://doi.org/10.1177/1098214016670029.

Vos, Leti, Michel Duckers, Cordula Wagner, and Godefridus van Merode. 2010. 'Applying the Quality Improvement Collaborative Method to Process Redesign: A Multiple Case Study'. *Implementation Science* 5 (1): 19. https://doi.org/10.1186/1748-5908-5-19.

Wells, Susan, Orly Tamir, Jonathon Gray, Dhevaksha Naidoo, Mark Bekhit, and Don Goldmann. 2018. 'Are Quality Improvement Collaboratives Effective? A Systematic Review'. *BMJ Qual Saf* 27 (3): 226–40.

Williams, N. J., C. Glisson, A. Hemmelgarn, and P. Green. 2017. 'Mechanisms of Change in the ARC Organizational Strategy: Increasing Mental Health Clinicians' EBP Adoption Through Improved Organizational Culture and Capacity.' *Administration and Policy in Mental Health* 44 (2): 269–83. https://doi.org/10.1007/s10488-016-0742-5.

Zammuto, Raymond F., and Edward J. O'Connor. 1992. 'Gaining Advanced Manufacturing Technologies' Benefits: The Roles of Organization Design and Culture'. *Academy of*

Management Review 17 (4): 701–28. https://doi.org/10.5465/amr.1992.4279062.