Research Article

Evaluation as a Multi-Ontological Endeavour: A Case from the English National Program for IT in Healthcare

Ela Klecun
London School of Economics
e.klecun@lse.ac.uk

Valentina Lichtner
University of Leeds
v.lichtner@leeds.ac.uk

Tony Cornford
London School of Economics
t.cornford@lse.ac.uk

Dimitra Petrakaki
University of Sussex
d.petrakaki@sussex.ac.uk

Abstract

This paper analyzes how researchers' different ontological and epistemological assumptions shape the process and outcomes of evaluation research. Focusing on the critical realism (CR) and social constructionism (SC) philosophical approaches, it outlines the rationale for multi-ontological evaluation and develops principles for conducting it. The paper draws from experience of evaluating a national implementation program of electronic health records in hospitals, one of the projects of the English National Programme for IT. It argues that an evaluation based on SC and one based on CR are significantly different in how they use knowledge gained in the field, and in the kind of evidence and recommendations that they can offer policy makers. The CR philosophy applied to evaluation research provides foundations from which judgments and abstractions can be presented in the form expected by the policy makers and managers who commission evaluations. In line with its ontological standing, social constructionism cannot simply or directly abstract and generalize across contexts, though it can offer other types of valuable evaluative insight. We show that, despite their differences, these two philosophical positions can, when taken together, produce jointly useful knowledge. This paper argues for the use of multi-ontological evaluation approaches and provides guidelines for undertaking such endeavors by emphasizing the need for mutual respect, dialogue, negotiation, and reflection.

Keywords: Evaluation, Information Systems, Ontology, Epistemology, Critical Realism, Social Constructionism, Health Care, Programs of Change.

Evaluation as a Multi-Ontological Endeavour: A Case from the English National Program for IT in Healthcare

1. Introduction

Information systems (IS) support and, at times, drive strategic policy initiatives, with complex technologies being implemented as part of large-scale programs of change (Dunleavy, Margetts, Bastow, & Tinkler, 2006; Fountain, 2001). The computerization of healthcare and the move to electronic health records (EHR) is perhaps the pre-eminent contemporary example of a major policy-led initiative that depends on the establishment of substantial IS and their committed use (Barrett, Oborn, Orlikowski, & Yates, 2012; Blumenthal, 2011a; Greenhalgh et al., 2009a; Greenhalgh, Potts, Wong, Bark, & Swinglehurst, 2009b; Oborn, Barrett, Davidson, 2011). In the past decade, most developed countries have pursued policies that endorse EHR systems as a central strategic goal for healthcare and have gone on to fund programs to promote and undertake the required work (AHRQ, 2010; Blumenthal, 2011a, 2011b; Brennan, 2005; Lang & Mertes, 2011; Léger, 2000). This commitment is based on expectations of achieving substantial benefits from EHR in various dimensions including economic efficiency, equality of access, care quality, and patient safety (Blumenthal, 2009; Blumenthal & Tavenner, 2010; Greenhalgh et al., 2008; Sheikh et al., 2011). These programs and the various projects they generate demand evaluation.

Evaluating IS (or rather the programs of change they are a part of) is a complex endeavor (Farbey, Land, & Targett 1999; Hirschheim & Smithson, 1988; Smithson & Hirschheim, 1998). IS implementations in healthcare are often large scale, costly, and complex infrastructure projects that involve or affect diverse and powerful stakeholder groups, that span long periods of time, and that engage different organizations and technologies (Ammenwerth, Graber, Herrmann, Burkle, & Konig, 2003; HoC, 2007; NAO, 2011; ONC, 2012). In the UK, as in many other countries, such projects have a high political profile (Sauer & Willcocks, 2007). The strong ethical traditions of medicine also add complexity to these projects as concerns for safety, privacy, and confidentiality are addressed. In recent years, medicine has also developed a strong commitment to evidence-based practice, which reinforces the need to evaluate IS.

Much evaluation activity around IS-led programs of change in healthcare is fundamentally political and is concerned with public accountability for resources used and the delivery mechanisms employed (AGoC, 2009; HoC, 2007; NAO, 2011; ONC, 2012). Some evaluations are narrowly economic and focused on cost and benefit, or financial sustainability (e.g., NORC, 2009). Traditional methods for these types of evaluations are "rooted in objectivist assumptions" (Friedman & Wyatt, 2006 p. 32) and built on a comparison of "outcome variables" across contexts or time, assessing performance by pre-established metrics, or focusing on a narrow range of anticipated effects. But, at other times, evaluation is more exploratory in purpose, and is expected to identify and account for wider changes occurring in the world as programs start to deliver new infrastructures and applications and as broader consequences (both intended and unintended) start to emerge (Ash et al., 2007; Greenhalgh et al., 2010b). This type of evaluation activity often takes the form of qualitative studies that focus on exploring processes of implementation and adoption and that emphasize the diverse interests of multiple stakeholders (Greenhalgh et al., 2010b; Takian, Petrakaki, Cornford, Sheikh, & Barber, 2011). This style of qualitative evaluation can serve to shape individual programs (e.g., as formative evaluation) and support wider processes of learning, helping to build an evidence base for IS-based innovations and thus to helping to steer longer-term policy.

Such "understanding-driven" evaluation (Hirschheim & Smithson, 1988; Smithson & Hirschheim, 1998; Smithson & Tsiavos, 2004) has often drawn on social constructionism (SC)[1] and emphasized the meanings people (socially) construct and ascribe to different phenomena; for example, the arrival of a particular IS seen as new technology, new managerial controls, or new work practices. This type of

---

[1] For the purpose of this paper, we adopt the expression "social constructionism" to represent our core understanding of what is a complex set of ideas where alternative nomenclatures, such as "constructivism", "constructionism", and "social constructionism", reflect diverse and nuanced views. In this, we follow Crotty (2003, p. 58, emphasis in original) and "reserve the term *constructivism* for the epistemological considerations focusing exclusively on 'the meaning-making activity of the individual mind'", while we use "*constructionism* where the focus includes 'the collective generation and [transmission] of meaning'". Thus "social constructionism emphasizes the hold our culture has on us: it shapes the way in which we see things…" (Crotty, 2003, p. 58). The IS literature, in contrast, has tended to use the blanket term "constructivism" when referring to those concepts. See also Hartwig (2007, p. 419).

evaluation with its SC roots highlights the socially constructed nature of evaluation itself (e.g., criteria and methods chosen) and of outcomes (e.g., expected benefits, judgments of "success" or "failure") (Lagsten, 2011; Wilson & Howcroft, 2005). Critical realism (CR), and more specifically the principles of realist (or realistic)[2] evaluation as developed by Pawson and Tilley (1997, 2004), has also been applied in evaluating programs of change, used in areas such as crime prevention and public health, and intended to support evidence-based policy making (Pawson, 2002). Recently, realist evaluation has gained influence in the health informatics community and in associated funding bodies because they seek to commission evaluations that will provide findings that are transferrable to other settings (Greenhalgh & Russell, 2010a; Oroviogoicoechea & Watson, 2009). In the IS literature, too, CR has been promoted as a basis for evaluation and as a way to provide robust generalizations that can transcend context (Carlsson, 2003; Dobson, Myles, & Jackson, 2007; Keller, Gäre, Edenius, & Lindblad, 2009; Mingers, 2004b; Wynn & Williams, 2012).

Both SC and CR are understanding-driven evaluations, where evaluators—their backgrounds, knowledge, and understanding of the subject of study—take a more prominent role. Evaluators are then key stakeholders in the evaluation, although often forgotten in projects' stakeholder maps. They have the hermeneutic task of disclosing and narrating the world for the commissioners of the study and for wider audiences. They must engage in reflective practice as part of their research method, and be aware of their biases and their "views of the world"—that is, their own philosophical assumptions (Alvesson & Skoldberg, 2000; Ashmore, 1989).

Debate on the implications of researchers' philosophical assumptions for studies of IS and for the IS discipline includes considering questions as to the validity and potential benefits of paradigmatic pluralism (Becker & Niehaves, 2007; Chen & Hirschheim, 2004; Mingers, 2001; Orlikowski & Baroudi, 1991). Similarly, philosophical assumptions underlying IS evaluations and the interests they serve have been discussed (Klecun & Cornford, 2005; Smithson & Tsiavos, 2004). The contribution of this paper lies in presenting a detailed and reflexive account of a study conducted by researchers belonging to different philosophical traditions. We draw on our own experiences as researchers, but this paper does not belong to a confessional genre, and it does not produce an autobiographical exposition, as done, for example, by Schultze (2000). Rather, we consider how our different ontological and epistemological assumptions shaped processes and outcomes of an evaluation of the English NHS Care Records Service (an EHR initiative of the English National Programme for IT— NPfIT). We assess, in particular, the potential of CR and SC as a philosophical basis for evaluation, and explore if and how working in tandem they may allow us to gain useful new knowledge about the phenoma being evaluated. Building on this, we outline the rationale for multi-ontological approaches to evaluation and develop principles for conducting such work.

The argument developed in this paper has five elements:

a) SC and CR each sensitize evaluators to different aspects under study and invite different evaluation questions

b) An evaluation based on SC and one based on CR are significantly different in their use of knowledge gained in the field, and in the kind of evidence and recommendations that they can provide for policy makers

c) Despite their differences, these two philosophical positions can, when taken together, produce complementary and jointly useful knowledge. In this sense they are compatible, although they are not interchangeable nor can they be simply synthesized

---

[2] Pawson and Tilley (1997) explain the word realistic (in realistic evaluation) as "a lovely compound word, plugging together the three Lego-bricks of *real… realist… realistic*" (p. xii, emphasis added)—"real*"* referring to the reality from an ontological perspective, *realist* referring to the methodology being scientific, and *realistic* because evaluation is "pursued in order to inform the thinking of policy makers, practitioners, program participants and public" (p. xiii), not for the sake of science. However, the expressions realist evaluation and realistic evaluation have been used interchangeably in the literature (e.g., Greenhalgh et al., 2009a).

d) Evaluators should reflect on their own philosophical positions and the role they play in their work. This is particularly true for non-positivist [3], understanding-driven evaluations; and

e) A dialogue or "trading zone" between different positions is useful for the outcomes of evaluations and for evaluators' processes of learning. Evaluations are fundamentally collective processes based on conversations and negotiations between stakeholders, including evaluators.

In presenting our argument, we contribute to the literature on the nature and practices of evaluation and to the discussion on multi-ontological research.

The paper is structured as follows. Section 2 outlines the ontological and epistemological assumptions of CR and SC and their implications for evaluation. Section 3 discusses the evaluation of NHS Care Records Service (NHS CRS). In Section 4, we develop two themes: evaluation as a process of judgement and learning, and evaluation as a multi-ontological endeavor. Section 5, the conclusion, discusses this paper's main contributions.

2. The Assumptions of Critical Realism and Social Constructionism and Their Implications for Evaluation

This section introduces the philosophical bases of CR and SC and explores how each may inform evaluations of IS implementation and adoption. The choice of these philosophical perspectives reflects the authors' individual beliefs and their research practice. It also builds on recent literature calling for a meaningful dialogue between these two perspectives (Elder-Vass, 2008; Mir & Watson, 2001; Newton, Deetz, & Reed, 2011). Our analysis first considers the ontological positions of both CR and SC, and then their epistemological assumptions (summarized in Tables 1 and 2, respectively, in Section 2.3).

2.1. Critical Realism's Ontological Assumptions

The cornerstone of CR is the coexistence of realist ontology and interpretivist epistemology (Sayer, 2000), which bridges the dualism between subjective and objective views of reality and avoids (what CR claims to be) the epistemic fallacy[4] of both interpretivism and empiricism (Mingers, 2004a). CR positions itself between positivism (or empiricism) and social constructionism (or post-modernism) (Mingers, 2004a; Pawson & Tilley, 1997; Potter & Lopez, 2001). Thus there is a real world that exists independently of our knowledge about it, although multiple interpretations of it may exist. Bhaskar (1978) proposes that this reality can be abstracted into three domains: the empirical, the actual, and the real (what we later refer to as the "EAR ontology"). The empirical is that which is experienced—events and things that are observed. The empirical domain is a subset of the actual—the domain populated by events (observed or not) and non-events (i.e., possible outcomes including those that have not occurred yet). The actual is, in turn, a subset of the real. The real also includes structures that give rise to generative mechanisms. The idea of a mechanism is central to CR ontology, reflecting the fundamental building block of causation that underlies the empirical and actual domains. As such, mechanisms represent the goal of much evaluation. It is the interaction of these mechanisms that causes the presence or absence of actual events. Confusingly, there is little agreement about what can be considered to be a "mechanism" (Horrocks, 2009; Wynn, Donald & Williams, 2008). Mechanisms have been conceptualized as causal powers, tendencies, ways of acting (Mingers, 2004a), emergent properties of certain relationships, or social practices (Horrocks, 2009). Examples of mechanisms reported in IS studies include: embededness of organizational routines, roles, and data in enterprise systems (ES) (Volkoff, Strong, & Elmes, 2007), characteristics

---

[3] Friedman and Wyatt (2006) distinguish evaluation methods between "approaches rooted in objectivist assumptions" and those "rooted in subjectivist assumptions". We struggled to find the right term that would include both SC and CR evaluations and do justice to their different philosophical positions, and therefore we used instead a negative notation with respect of the philosophy they both subscribe to ("non-positivists"). In doing this, we do recognize that CR has some positivist elements, but, as explained in Sections 2.1 and 2.3, the stratified ontology is hardly "objectivist", and the epistemological stance is clearly interpretivist.

[4] The epistemic fallacy lies in the reduction of the ontological domain to the epistemological domain.

of ES and organizations as an explanation of organization-ES misfits (Strong & Volkoff, 2010), and innovation and service mechanisms (Bygstad, 2010).

CR ontology is explicit in encompassing both natural and social worlds. However, it acknowledges that, unlike the natural world, the social world depends on human action for its existence. This does not mean that each individual constructs anew their own social world. On the contrary, CR suggests that we come to live in a world that is pre-constructed and that has emerged from human activities but is distinct from them. As such, the world cannot be reduced to contingent enactments and local practices, which a strong social constructionist viewpoint might suggest. There are pre-existing structures that are the result of past actions. These (social and material) structures both enable and constrain activities, such as organizational action and communication, and are reproduced or transformed by such activities. These structures have emergent properties and cannot simply be reduced to their constituting elements (Elder-Vass, 2008). Hence, for CR, structure and agency (individuals with their capacities for agency) should be understood as analytically distinct (Fairclough, 2005; Reed, 2000). For example, individuals and their practices should be seen apart from (although interlinked with) social forms (Archer, 1998); the role of, for instance, a nurse and what it entails (structure, historically handed down) is analytically different to a particular person who works as a nurse and in their own way fulfils that role.

Technology is conceived of as "material"—a structure—and as having properties or features (Mutch, 2002, 2010; Volkoff et al., 2007; Wynn, Donald & Williams, 2008). It cannot be reduced to users' interpretations, or to the way it is filtered through practices. Conceptualizing technology in this manner allows researchers to explore how they evaluate the interrelationship of technologies and their features when set in different contexts (Mutch, 2010). Technology can render some aspects of structures more durable in time and space (Mutch, 2010) when organizational routines, roles, data, and the relationships between them become embedded in technology and hence acquire a material and durable aspect (Volkoff et al., 2007). For example, EHR software as a durable structure can enforce a separation between what a nurse can do and what a doctor can do, one that may have been more malleable in the previous paper-based system. Conversely, the choice to develop a particular technology and its constitution (shape and type of features) is influenced by existing structures and their emergent properties (Mutch, 2010).

For social constructionists, a cursory read of CR principles as outlined above might result in a belief that CR is technologically deterministic, particularly in that objects are seen as potentially having causal powers. However, the notion of causality as being contingent means that placing the same technology in a different context does not imply that the same mechanisms will be activated (Smith, 2006). Thus, CR does not suggest that technology simply "impacts" on society or an organization in a single and pre-determined way, but rather that technology can be seen as having potentiality for consequences.

2.2. Social Constructionism's Ontological Assumptions

The view of SC adopted here, in common with most works in IS, asserts that "all knowledge, and therefore all meaningful reality as such, is contingent upon human practices, being constructed in and out of interaction between human beings and their world, and developed and transmitted within an essentially social context" (Crotty, 2003, p. 42) (see also the discussion in the introduction to Latour's *Pandora's hope* (1999)). In other words, for SC, it is not that the world does not exist outside our mind, but that its meaning (i.e., the identity of a thing, rather than its interpretation) is ascribed by our consciousness. These meanings are generated collectively (socially) rather than independently by each individual (i.e., meanings "precede us" and "we are already embedded" in them (Crotty, 2003, p. 52)). Human beings engage with the world that is "always already there" (a view espoused by both Heidegger and Meleau-Ponty). This emphasizes the historical and situated aspects of meaning generation and transmission—a position not so far removed from CR. However, in SC, there are no structures as such "out there" and independent of our perceptions. If the term "structure" is used in SC, it tends to be understood as intersubjective and as constructed through lived experience. For instance, constraints of roles (e.g., that of a doctor, a manager, or an evaluator) arise not from the role by itself, but from the intersubjectively created notion of what the

role is (Willmott, 2005). Boundaries and the properties of "concepts" and "things" only emerge "from their mutual intra-actions" (Barad, 2003, p. 820). Therefore, materiality and use (or function and meaning of objects) are "not distinct moments, but emerge in an on-going materializing performance of the world" (Iedema, 2007, p. 941). This represents a move toward a performative understanding of the world and a concern with practices, doings, and actions (Barad, 2003, p. 802). The reality is, or rather realities are, constructed in provisional, ongoing, and dynamic acts. They are products of different contexts, perspectives, and sense-making mechanisms (Guba & Lincoln, 1994; Mir & Watson, 2001, p. 1172). Gergen (1994) states that "there is no foundational description to be made about an "out there" as opposed to an "in here", ...once we attempt to articulate "what there is"...we enter the world of discourse" (Gergen, 1994, p. 72, cited in Nightingale & Cromby, 2002, p. 703).

Growing interest in such accounts in IS has given rise to studies showing that the boundaries of what is considered the technical and the social are fluid and re-negotiated in practice (Bloomfield & Vurdubakis, 1994), and that technologies "acquire form, attributes, and capabilities through their interpenetration" (Orlikowski & Scott, 2008, p. 455). In organization studies, a similar shift from an emphasis on organizations as structures to organizations as processes, expressed in ideas of "organizing" or "organizational becoming", is occurring (Tsoukas & Chia, 2002). The centrality of discourse in creating organizational reality is also being asserted—a move strongly opposed by critical realists such as Reed (2004) and Fairclough (2005).

2.3. Critical Realism and Social Constructionism's Epistemological Assumptions

For the practice of evaluation, epistemological questions are central. Evaluation of any kind has to negotiate access to the object of study and justify itself in terms of the fidelity of the evidence transmitted from the field in question.

CR proposes that our knowledge about the world is limited and fallible. Reality cannot and should not be equated with the inadequate and partial knowledge we obtain. However, the interpretivist epistemology of CR does not imply that all interpretations represent reality equally well. Indeed, CR evaluation has, as a central goal, to establish whether some representations are better holders of knowledge about the world than others (Fairclough, 2005), and thus places social scientists or evaluators in the role of arbitrator or judge (Newton et al., 2011)[5]. This is a position disputed by SC, which sees production of knowledge as context dependent, inevitably embedded in social and political processes. Judgment of the "subject matter" cannot be made because it is not possible to know the "subject matter" independently of the discursive practices that are invoked to disclose it (Willmott, 2005, p. 762).

Figures 1 and 2 present a (simplified and stylized) visualization of ontological and epistemological assumptions of CR and SC, highlighting the role of a researcher.

---

[5] However, more recently, Al-Amoudi and Wilmott (2011), considering potential convergences between SC and CR, suggest that epistemological relativism could be seen as central to CR. In the words of Potter and Lopez, "critical realism is a broad church" (2001, p.15), and we cannot claim to represent all views in the limited scope of this paper.
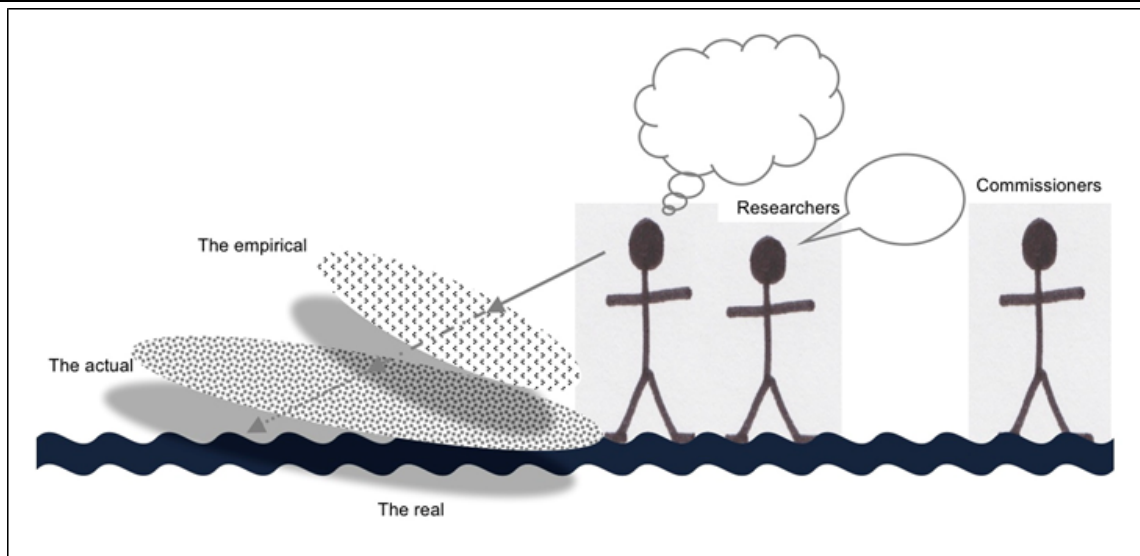
Researchers, trying to generalize to the actual and making inferences on the real (one a subset of the other), observe and interpret the empirical; they report back (on outcomes, mechanisms, and structures) to the commisioners.

Figure 1. CR Ontology and Epistemology
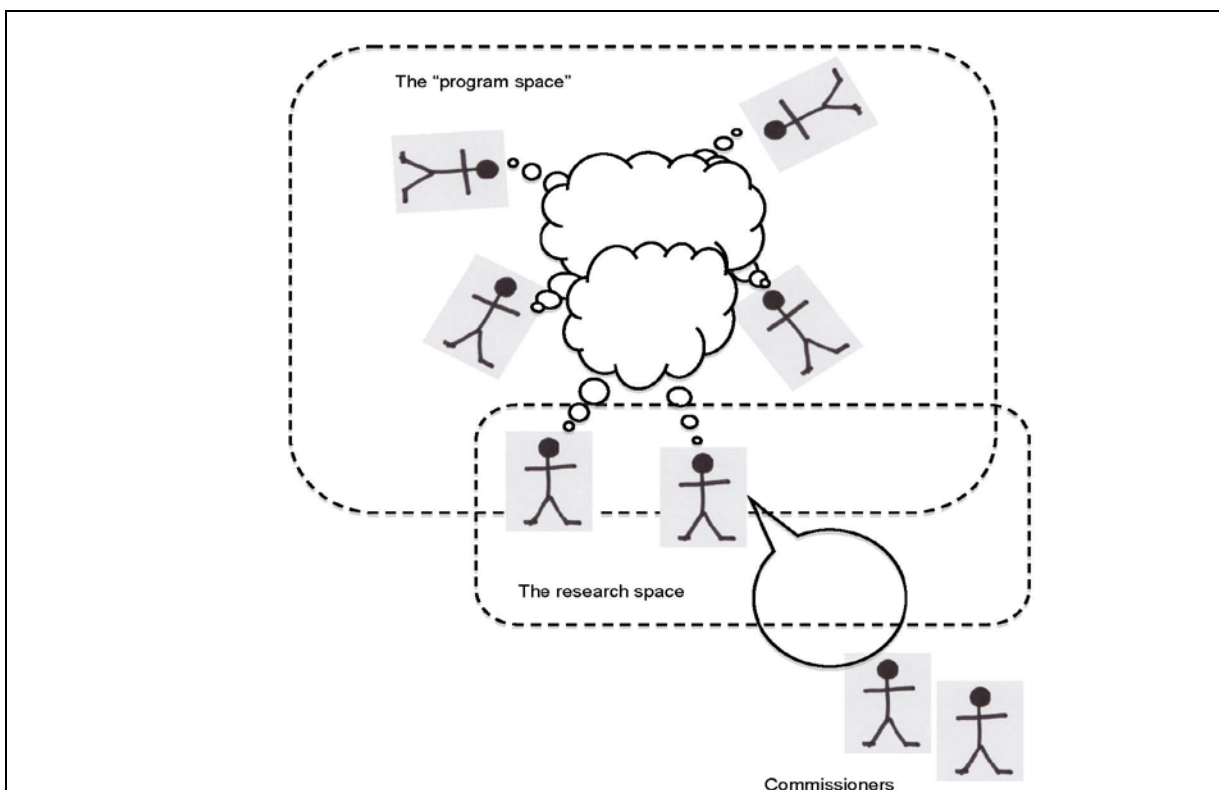
People engaged in the IS program of change (IS implementation, healthcare work) co-construct their reality. Researchers approach the object of study and come to share it, and therefore themselves co-construct it. They convey their interpretation of this to the commissioners.

Figure 2. SC Ontology and Epistemology

The ontological and epistemological assumptions discussed in Sections 2.1-2.3 imply different understandings of what evaluation is and how it should be conducted. Specifically, differences arise with regard to conceptualizing what is to be evaluated (e.g., how technology is understood), evaluation goals, or, more generally, the type of knowledge sought, and the interpretation and reporting of this knowledge. Tables 1 and 2 below summarize the main assumptions of CR and SC and their implications for evaluation.

Table 1. Main Tenets of Critical Realism and Their Implication for Evaluation		
	Main tenets	**Implications for evaluation**
Ontology	Real world exists independently of our knowledge about it.	Evaluation intends to reveal the reality (or rather produce best approximation of it).
	Stratified ontology: real, actual, and empirical.	Evaluation aims to observe empirical events, form hypotheses concerning the causal powers that produce them and identify and explain the mechanisms that support such causal powers.
	Structures and individuals are understood as analytically distinct.	Evaluation should seek to describe how structures enable and constrain activities.
	Technology has inherent properties or features that cannot be reduced to their interpretations by users or to the way they are used.	Evaluation should explain technology in terms of its properties or features and how these are interpreted and used.
Epistemology	We cannot equate reality with our knowledge of it, which may be limited and fallible.	Evaluation is never "perfect". There are always some things that may be undiscovered or misunderstood.
	Some representations of the reality are closer to the real than others. Evaluators should provide reliable representations.	Evaluation needs to seek to produce the best representation possible. Evaluators need to judge representations according to their reliability.

Table 2. Main Tenets of Social Constructionism and their Implication for Evaluation		
	Main tenets	**Implications for evaluation**
Ontology	There is no one "objective" reality but many realities produced and re-produced by people.	Evaluation cannot "reveal" reality. The process and outcomes of evaluating co-construct reality.
	The world is fluid and undetermined.	Evaluation explores and describes how boundaries between objects/concepts are constructed and maintained.
	Structures, if any exist, are not independent of individuals; individuals enact and (re-)produce structures.	Evaluation needs to focus on how realities are brought into being, through interpretations and enactments.
	Technology does not have inherent properties but is socially constructed (by designers, business analysts, users, managers, etc) and enacted.	Evaluation explores technology-in-use (i.e., technology as it is performed through practices).
Epistemology	Production of knowledge is context-dependent, embedded in social and political processes.	There cannot be truly "independent" evaluation and "objective" evaluator positioned outside the subject of evaluation.

2.4. Generalizability of Findings: Causality and Generative Mechanisms

Any evaluation—and especially an evaluation of a large-scale program of change across multiple organizational settings—must address issues of representativeness and generalizability. Indeed, issues of generalizability are core to both the IS field (Lee & Baskerville, 2003) and the health informatics field (Scott, Prytherch, & Briggs, 2010). Both IS and health informatics consider themselves as applied disciplines that aim to inform practice and policy, which, in turn, implies that evaluation will draw and present conclusions as some form of judgement.

For CR, generalization is not based on uncovering universal laws; instead, it relies on identifying an "underlying essence of things" (i.e., generative mechanisms, structures, and the relations between them) (Danermark, Ekstrom, Jokobsen, & Karlsson, 2003). Causality may also include accounts of intentions and feelings (Mingers, 2004b). From a SC perspective, this formulation of generalizability sounds suspiciously essentialist despite CR's assertion that generative (causal) mechanisms are tendencies or possibilities that can be actualized or not. Mechanisms, it is argued, operate in a highly contingent, contextualized and less-than-deterministic manner. Most SC proponents would consider the abstraction of essential features and identification of generative mechanisms as impossible exactly because they are context generated (e.g., as a technology-in-practice (Orlikowski, 2000)), and hence, in abstraction, they become meaningless. Put in stronger terms, we might say that, ontologically, generative mechanisms are non-existent, and, epistemologically, they are a matter of interpretation.

Critical realists believe that, although knowledge is fallible, limited generalizability is possible through accumulating knowledge over time (e.g., from different case studies) and applying systematic scientific research methods. Through the process of testing hypotheses and eliminating alternatives, evaluation researchers can engage in a process of theory building (e.g., a theory of IS implementation in organizations) (Carlsson, 2011; Pawson & Tilley, 1997).

SC studies, in contrast, do not usually set out to produce normative scientific claims or generalizable results (in the usual sense). Research objects, process, and outcomes are all seen as being subjective (Weber, 1949). Yet, studies that claim to be social constructionist, including those reporting on evaluations of IS implementations (e.g., Jones, 2003), often suggest some "lessons to be learned". From a CR perspective, these are equivalent to generalizations beyond the context of study. SC researchers though may claim that these constitute statements that describe how the object of research tends to be (i.e., an ideal type rather than how it should be). They function as a point of comparison without, however, being a normative model (Weber, 1949).

3. The Evaluation of NHS Care Record Service

3.1. The Project

In September 2011, the U.K. Department of Health announced the "dismantling [of] the NHS National Programme for IT"[6]—the biggest (non-military) IT program in the world (Brennan, 2005) launched in 2002 to computerize and "modernize" the English NHS through a variety of IT implementation projects and change programs (NAO, 2011). A large part of the initial £12bn (US$18bn at the time of writing) NPfIT budget was to be spent on the implementation of a detailed EHR for all hospital and community health settings under an NPfIT initiative known as the Care Records Service (NHS CRS). NHS CRS was based on a limited number of EHR commercial software applications including iSOFT Lorenzo and Cerner Millennium for secondary care (hospitals) and CSE Servelec Rio for community mental health[7]. These systems were pre-selected for specific geographical areas (also known as clusters) (Table 3). Roll-out of these systems started in the mid-late 2000s and proceeded across England. The leading role of driving the implementation of NPfIT projects and systems was given to

---

[6] Press release: http://mediacentre.dh.gov.uk/2011/09/22/dismantling-the-nhs-national-programme-for-it/

[7] Health care provision in England is divided into two main areas: Primary Care and Secondary/Tertiary Care services. Hospitals are part of the latter. Primary care services include NHS Direct, GP practices (family doctors), NHS dentists, community pharmacists, and opticians. Mental health services are seen as a specific area, with care provided in the community and/or in hospital (see http://www.nhs.uk/NHSEngland/AboutNHSservices/Pages/NHSServices.aspx)

Connecting for Health (CFH), which was initially set up as an independent entity. During the time of this study, the responsibility for NPfIT was taken back into the direct control of the Department of Health. At the outset (2002), consortia of private companies designated as local service providers (LSPs) were contracted to undertake the implementation and ongoing support of these systems on a regional basis. The delivery of care by the NHS was organized into trusts[8], and these have become, in recent years, increasingly financially independent ("foundation trusts"). Thus, the implementation of these top-down centrally selected systems depended also on choices made locally by individual NHS trusts.

Table 3. NHS CRS Delivery Structure for Secondary Care (as of March 2011)					
	Clusters				
	North, Midlands and Eastern (NME)	**London**		**Southern**	
Local service provider (LSPs)	CSC Alliance	BT		No local service provider*	
NHS service	*Community & hospitals*	*Community*	*Hospitals*	*Community**	*Hospitals**	*Community & hospitals*
Software supplier	iSoft	CSE Health care (formerly CSE-Servelec)	Cerner	CSE Health care (formerly CSE-Servelec)	Cerner	Various suppliers with CFH Additional Supply Capability and Capacity framework contracts
Software application	Lorenzo Regional Care	Rio	Cerner Millennium	Rio	Cerner Millennium	Various

Note: *After the departure of Fujitsu as LSP for the Southern region, BT was contracted to take over a limited number of Cerner Millennium and Rio implementations.

NPfIT was accompanied by a program of independent evaluations, the Connecting for Health Evaluation Programme (CFHEP), run by the University of Birmingham[9]. Among these evaluations was the "Evaluation of the Implementation and Adoption of the National Health Service Care Records Service", which we draw on here.

This evaluation was commissioned in late 2007, started at the end of 2008, and run for about two-and-a-half years until March 2011, when the final report was published (Cresswell et al., 2011; Robertson et al., 2010; Sheikh et al., 2011). The project is an example of multidisciplinary research (advocated, for example, by Payton, Pare, Le Rouge, and Reddy (2011)). It was undertaken by a team of researchers from four U.K. universities, who represented different disciplines, including medicine, pharmacy, health management, information systems, human computer interaction, and social sciences. As individual experts, our ontological and epistemological stances to evaluation could

---

[8] See NHS Choices website—Authorities and trusts: http://www.nhs.uk/NHSEngland/thenhs/about/Pages/authoritiesandtrusts.aspx
[9] See http://www.haps2.bham.ac.uk/publichealth/cfhep/

be mapped along the entire spectrum—from "positivist"[10] to SC, including CR somewhere in between. Our multiple perspectives influenced how we conducted the evaluation, the evaluation's research design and methodology, and the final report's analysis and writing. Our understanding of our primary target audience's expectations—the policy makers of the Department of Health—also influenced the final report's analysis and writing. However, we also tried to satisfy other audiences such as academics (Friedman & Wyatt, 2006).

3.2. Research Design

The evaluation's commissioning brief expected a "systematic" evaluation of NHS CRS systems, studied at "an organizational level", with the intent "to determine how organizational dynamics affect the uptake of the service", and assessing "its impact on patient care in England" (University of Birmingham, 2007). Although both qualitative and quantitative methods were explicitly welcomed, the examples provided in the commissioning brief all assumed a positivist view of the world, reflected in the assumption that statistical analysis of the results would be conducted, as the following extract shows:

> *For example, applicants could evaluate the development of the NHS CRS in hospitals using a naturalistic (non-experimental) cluster stepped wedge design, where institutions are recruited at different stages of the implementation of the NHS CRS and then tracked over a number of time periods. Alternatively, the assessment could be carried out by comparing sites randomly selected from the group of early-adopters with a random selection of sites yet to implement the function. Studying a large number of sites will improve statistical precision in the conventional manner and will also allow variation between implementing and non-implementing sites to be estimated much better.* (University of Birmingham, 2007)

In answering the call, our research protocol proposed both formative and summative evaluation guided by realistic evaluation tenets (Pawson & Tilley, 1997) and Cornford, Doukidis, and Forster's (1994) sociotechnical evaluation framework. Thus, a key research question was: "Which mechanisms of the NHS CRS work for whom under what circumstances?". Our research design did not explicitly propose a critical approach, which is in line with most realist evaluations where the critical dimension of CR is underplayed (Carlsson, 2003; Pawson & Tilley, 1997; Porter & O'Halloran, 2012).

Our evaluation was to be prospective, longitudinal, and real-time, and conducted while NHS CRS implementations were expected to be taking place. Selection of and access to sites was challenged in practice by implementation delays of NHS CRS systems (the logistical and methodological challenges this caused are discussed more in depth in Takian et al. (2011)). Thus, we replaced the original "step-wedged" design for systematic sampling of NHS trusts with purposive identification of trusts of different status (e.g., hospitals versus community mental health settings), geographical implementation areas (i.e., North Midlands and East, London, and Southern England) and choice of NHS CRS software system.

In the course of our research, partly due to these circumstances, and partly due to fundamental ontological shifts experienced by some of the researchers won over by SC arguments, the object and purpose of the research changed from assessing what the NHS CRS implementation and adoption achieved (and to what extent the initial goals were met) to an understanding of NHS CRS "in-the-making". This process of shifting and a journey of discovering the SC perspective undertaken by some of the evaluators is reflectively described in another paper (Takian et al., 2011).

Seventeen NHS care-giving organizations were involved in our study, of which 12 were treated as in-depth "case studies" (the sites are identified by a random letter; e.g., "case R" or "case P"). We approached the other five trusts for recruitment but they did not give us sufficient access to conduct in-depth and longitudinal studies; what we learnt of their experiences informed our cross-site analysis, but they did not constitute case studies of their own. The organizations we studied were among the

---

[10] Since this paper focuses on CR and SC perspectives, we do not analyze or discuss those study elements that were influenced more by the positivistic approach.

first to receive (or were waiting for) NHS CRS software systems as part of the national program. We negotiated for research access locally with key contacts and gatekeepers, and we matched local preferences for recruiting participants with the suitability of research methods (e.g., a survey method was only accepted in one of the 12 NHS trusts). Most of the interviews and observations took place between February 2009 and January 2011, and usually over two periods of time in each site, with a six to nine months gap.

At the end of our study, information for our analysis was drawn from a total of: 431 semi-structured interviews (face-to-face, by telephone, or email); 590 hours of observations; 234 sets of notes from observations, researcher field notes and conferences; 809 NHS documents; and 58 national and regional documents. In the majority of cases, interviews were audio-recorded and transcribed, and, when research participants preferred not to be recorded, field notes were taken. Individual researchers were granted access to conduct non-participant observation in study sites: they shadowed users of the NHS CRS systems and attended meetings.

Among the stakeholders interviewed or observed were clinicians (41%), including physicians, nurses and allied health professionals; managers and administrative staff (26%); IT implementation staff (18%); representatives from NHS Connecting for Health (8%) and local service providers (2%); and patients and carers (2%) (see also Sheikh et al., 2011). We also collected public NHS trust documents and, with permission, confidential ones. These included deployment timelines for the NHS CRS, project initiation documents, business cases, risk registers, minutes from board meetings, "lessons learnt" documents, training strategy documents, and annual reports.

Because the evaluation was intended to explore NHS CRS implementation and adoption at an organizational level, each organization was treated as a case to be studied in its own right. Designated researchers in our team led each specific case study and their disciplinary background and theoretical stance influenced the way each case study was undertaken, as reflected in the "data"[11] collection and "data" analysis strategies. Section 3.3 discusses how these different theoretical stances reflected our individual and collective approach to research, analysis, and writing of our "findings".

3.3. Researching, Analyzing, and Writing up the "Findings"

Researchers in our team with a CR background carried out interviews as "fact-finding" missions to inform the drawing of evaluative conclusions and to hear people's views of events and NHS CRS systems. They considered NHS CRS systems as artefacts that could be "objectively" identified, interpreted, and compared in their technical design, functionalities, and qualities. In contrast, researchers with a SC background perceived the evaluation process as an opportunity to explore meanings, what NHS CRS meant to the people involved in the implementation process, and how they "constructed" it in their own context. They approached NHS CRS as an object in the making rather than a given system to be evaluated. For them, there was no "a priori data" to draw on for evaluation. Ontologically, the study was not concerned with an "objective" technology but with the "biography" (Williams & Pollock, 2011) of a technology, when feasible, and with a "technology-in-practice" (Orlikowski, 2000).

To help develop a shared (if tentative) understanding of what and how we should evaluate, we conducted regular multi-disciplinary qualitative workshops among the researchers. We started by questioning the assumptions behind the commissioning brief and our initial research protocol. We argued about our research design, methods, and the nature of what we were to study. Implicitly and explicitly we questioned each other's ontological and epistemological assumptions. Through on-going dialogue we created a shared (methodological) language that enabled us to work out a single research design while accommodating different ontological perspectives.

Our team's SC and CR researchers shared a belief that each case should be understood in its entirety (in its context) and in depth. To produce the final report and meet our obligations to the

---

[11] In this paper, we use the terms "data" and "findings" with reluctance and we place them in quotation marks. Both terms imply ontological assumptions, acceptable for CR, but incompatible with a SC perspective.

sponsors, some analysis across sites was required to elicit main themes and construct lessons learned. We agreed on a set of main themes and then relied on each researcher's own interpretation of their cases—each one selecting the significant ideas or features, supported by quotes/field notes in relation to each theme. Each researcher was given a section of the final report (e.g., a theme) to write and was asked to summarize their colleagues' analysis while trying to do some justice to everybody's interpretation. We all then reviewed and commented on the entire report with the opportunity of editing sections that did not represent our views, although for some of us this necessarily required some compromise.

In Sections 3.3.1.and 3.3.2, we present examples of how site-specific analysis reflected SC and/or CR understandings. We also discuss how, while a common, more-or-less uniform thread was achieved in the cross-cases analysis, still these different understandings remained visible in different chapters of the report. The quotes are from our report.

3.3.1. Site-Specific Analysis

3.3.1.1. Vol.2, Chapter 13: "An Example of Self-Evaluation in Site R"

One of our case studies was "site R". The trust in "site R" had implemented one of the NHS CRS systems but, about 2 years from going live, had decided to "switch the system off" and return to a version of another software application that they had used previously. A SC researcher led the study of this case and interviewed managers and clinicians on site. Her aim, as she reported in the case study, was to:

> *explore the way in which* [the system] *was introduced and implemented in Site R, to capture participants' initial expectations and experiences of the implementation process, and to collect their reflections on the processes they followed and on the things they would like to have done differently.*

The interviews covered a variety of themes—from the visions and rationales for undertaking the implementation, to the implementation strategies, the process of software configuration, training and organizational learning, the use of the software, and the shaping of users' experiences. In terms of "system in use", in the eyes of the people who had struggled with the system for more than 12 months, the application was not fit for purpose: its interaction and interface design was cumbersome; functionalities were lacking. For example, a clinician recounted a problem with the length of data fields:

> *If it said myocardium infarct suspected, it would just say myocardium infarct; or myocardium infarct ruled out and it would say myocardium infarct because it only had that space. Amazingly, how could you design a clinical information system that didn't even have truncation sorted out? Just unbelievably basic.* (Interview, healthcare professional)

A CR researcher would have made a value judgement. They would have seen this as an iconic example of a "wrong" top-down approach that had imposed a "badly designed" software application that could not be made to work effectively—a fundamentally faulty mechanism. The CR researcher might then have tracked the ultimate generative causes of this "disaster"—that, as a participant described, "has cost us millions of pounds. It's brought our hospital nearly to its knees" (interview, healthcare professional)—to the design of the technology. From a CR perspective, this was a case of the faulty inherent properties of the technology, which were ontologically separate from interviewees' interpretations (who either confirmed or disconfirmed). Value judgments of the technology could thus be made as long as sufficient evidence of the empirical was found.

In contrast, the SC researcher's conclusions were not judgemental. She presented the study as a case of "self-evaluation"—she left the participants to evaluate and judge their own experience. According to her, value judgments of the technology and its implementation could only be made by exploring participants' own interpretations and accounts of the technology-in-use. She saw the case of site R as *"an iconic example of the complexities that surround implementation of EHR in hospitals

and [as] indicative of the sociotechnical implications that such endeavors have". She believed that causes of success and failure could not be found "out there", and she saw them as being ontologically subjective. She thus intended to bring together conditions of potential success and failure by exploring participants' constructions. As a result, generic conclusions, and not value judgments, could be arrived at by making reference to "the way in which such IT initiatives are managed" and "the impossibility of rendering technology meaningful in a context whose attributes are not reflected and cannot be inscribed into the technology".

3.3.1.2. Vol. 2, Chapter 14: "The Evolution of Electronic Health Records in the South"

Short studies carried out in three NHS trusts and two county-wide health informatics services authorities (HIS) in the Southern cluster were reported together in one chapter of the report. Research in two NHS trusts was led by a SC researcher, while the other three sites (one trust and 2 HIS) were studied by a CR researcher. While the researchers jointly wrote the chapter, the SC researcher took the lead in writing up the case and wrote the concluding remarks. These she identified as:

> *Three main conditions which framed and help to explain the evolution of the NHS CRS in the South*:
>
> 1. *The regional management of the programme*
>
> 2. *The level of configurability of the software,* [and]
>
> 3. *The context within which the NHS CRS initiatives were carried out.*

The South had experienced a different NPfIT story than the rest of England. As the LSP responsible for implementing NHS CRS in the South had left NPfIT, local organizations had the choice of implementing an EHR system and/or the NHS CRS "national vision" in their own way. The two local NHS HIS studied in the South were engaged in the early stages of implementing their version of NHS CRS by building portals that would enable hospitals, community organizations, and primary care organizations to share patient records. Thus, in their eyes, they were "pursuing the NHS CRS dream" (field notes, senior manager) by integrating NHS care through integrated systems.

In the meantime, in nearby counties, the three NHS trusts studied were moving their own implementation of software systems forward: one with a different version of Cerner Millennium (Site F), one (a mental health trust) moving from Millennium to Rio (Site BB), and one by building an EHR by way of adding software modules to their existing patient administration system (Site P). Most NHS trusts in the South had suffered uncertainties over the NHS CRS program after the departure of the LSP. All researchers agreed that such aspects of the wider context in which NHS CRS initiatives were carried out greatly influenced the implementation and adoption processes. These three sites had each eventually managed to find their own solutions to the implementation. The SC researcher found an important aspect of the millennium implementation in Site F, which was the trust's intention to change the organization while adapting the system to the organization's needs:

> *The purpose* […] *was neither to adapt the system to what they were doing nor to change what they were doing in order to fit the system. Rather, they intended to achieve synergies and innovate by making mutual adaptations both to their work practices and processes and to the system.*

The SC view was highlighted in this theme, which reflects a belief that implementation and adoption are multidimensional and dynamic processes of social construction, where both technology and people adapt and change.

A similar process of adaptation was also seen in Site BB—a trust that had tried to implement Millennium but had found this unsuitable for mental health and therefore moved to Rio. One of the most interesting themes, from both SC and CR perspectives, was the nature of work in this setting

(mental health): its geographical distribution across hospitals, community settings, and patients' homes, and the narrative form of clinical notes:

> *I would think that actually your computerized system is even more important for mental health service than hospital, because of the very nature of how we work. The sense is that we are scattered about and we are based in one place and we see our patients in another place and we see patients at home and patients in A&E department somewhere.* (Healthcare professional, Site BB)

> *…we would write notes and so on. Notes would go into pages and pages, sometimes on each assessment.* (Healthcare professional, Site BB)

Both SC and CR researchers agreed that the mental health context was fundamentally formative for understanding the implementation choices made in this trust. From a SC perspective, these choices were not presented in terms of causal mechanisms that determined the move from Millennium to Rio, but as a result of a process of learning and adaptation, which was conditioned on the way in which clinicians perceived technology and their work—its nature, practices, and habits.

Two CR generative mechanisms that were in the research analysis, but eventually not included in this case study by the SC researcher leading this study, were:

- Challenges of information governance and interoperability standards required for systems to share data across settings, and

- NHS structural arrangements that placed trusts in competition with each other.

For the CR researcher, these were seen as mechanisms undermining the NPfIT vision of the NHS CRS which emphasized the sharing of patient records across healthcare providers. Implicit in this is the assumption that structures closer to the NPfIT vision would have led to a successful Millennium implementation. While interoperability could eventually be agreed on among stakeholders, the NHS structure and competition arrangements seemed to be stronger mechanisms that worked to undercut such ambitions for CRS. From a strategic point of view, for the NHS as a whole, integrated and centrally driven IT projects were seen as enablers for an essential re-organization of the NHS supporting the creation of "specialized centers in the region, optimizing costs, serving the larger population of the region, since they can all access local information" (Regional senior manager, March 2009). However, the local HIS struggled to achieve collaboration among local NHS trusts and primary care trusts (NHS commissioners) towards sharing data.

> *We are aligned with the PCT's[12] transformation strategy but it is difficult to engage with people involved in strategic planning. How are they going to support the exchange of information* […]? […] *people are acting very parochially, thinking only in terms of their problems and solutions.* (Senior manager, Jan 2010)

Contrary to the strategy being pursued by this HIS, NHS hospital trusts saw NHS CRS systems as software necessary for their own survival in a competitive market with a reduced overall budget: "The electronic health record was seen as a necessity that would help to improve efficiency and ensure the future survival of the trust" (Site F). The IT manager of a local hospital (Site P) interpreted the HIS portal project as providing benefits more to the community than to the trust. He explained that the trust faced limited competition in its region and it had no interest in enhancing competition by contributing to an integrated IT system. The CR researcher analysis went beyond the organizational-level and considered structural characteristics in the marketplace undermining the potential to achieve the policy vision, and deriving recommendations for change (e.g., a recommendation for alignment between NHS funding policy and NHS IT strategy). SC researchers did not see the

---

[12] Primary care trusts (PCTs)—NHS organizations responsible for developing health services for a particular community (now abolished).

"marketplace" as an external structure, but rather as a social construct produced by interpretations of national policies and local strategies by key stakeholders. This opened up opportunities for different stakeholders to pursue their interests.

3.3.2. Cross-Site Analysis

Most of the themes that emerged from the case studies made it into the final cross-site analysis and are included in the volume 1 of the report. An important theme that emerged from a number of sites analyzed from a SC perspective related to the multiplicity of visions associated with NHS CRS:

> *4.4.2 What is the NHS CRS?*
> *Our findings have indicated that the NHS CRS was not only the generic name given to a range of software systems (i.e., Lorenzo, Millennium and RiO), but also a multi-faceted concept. Different research participants (i.e., managers, health care professionals, administrative staff and patients) attributed a different meaning to the nature and role of the NHS CRS in secondary health care settings.* (Vol. 1, p. 75)

All researchers were invited to contribute to this section of the report with quotes and field notes that would support this theme. The SC researcher leading the writing of this section presented the different meanings people attributed to NHS CRS without any reference to "an official vision" (e.g., the national vision, or what was originally envisioned for NHS CRS at the start of NPfIT). Rather, she listed them by introducing each one with clauses such as: "The NHS CRS was conceptualized by participants…" or "interviewees described the NHS CRS as…" or "In contrast to this, some participants saw the NHS CRS as…". It was made clear that some views were contradictory, but there was no attempt to assess them against each other. Furthermore, no attempt was made to find an underlying direct cause for the different views. Instead, focus was placed on conditions that encouraged certain interpretations over others, conditions that were also recounted by participants rather than discovered deliberately or analytically by the researcher. The rationales were those of the interviewees, their interpretation of the causes (e.g., because they saw a potential in NHS CRS to improve old-fashioned ways of working). For instance:

> *The NHS CRS was also presented as an indication of modernization, a necessity and "a way of the future" (Interview, Patient, Site H). This was not only because of the availability of advanced technology, but also because of its potential to improve the old-fashioned ways in which the NHS works … In contrast to this, participants from Site M perceived the NHS CRS not as a means to change, but as a way to reinforce already existing policies…* (Vol. 1, p. 78)

Those different interpretations were of importance in themselves to SC researchers. It is through exploring those interpretations and how they have evolved over time that a better understanding of different approaches to implementation and use could be achieved.

Had there been space in the already extensive report for expanding this theme, the CR researcher might have presented a different analysis of these multiple visions by assessing the value and feasibility of the national vision (e.g., as presented in policy papers), then assessing in what way this vision had been translated in local enactments, and whether contradictory visions were to be understood as mechanisms reflected in problematic implementations. She would then have considered reasons for those translations and potential "distortions".

In SC language, interpretations and visions were part of a dynamic process of change—"a process that was made to work through the mediation of a number of people and technologies" and constructed "primarily on relationships, processes and practices as they emerged on the surface" (Vol. 1, p. 97). As new constructions and relations built on the old, this process of change "is then a process surrounded by continuities and discontinuities of ways of acting and thinking" (Vol. 1, p. 97).

The CR researcher also saw implementation as a dynamic process, but her approach to understanding it differed from the SC approach. First, in describing how processes of NHS CRS implementation and

adoption were blurred, the CR researcher used people's quotes to describe what she saw as facts (e.g., "Every time you get a new build you sometimes get some regression...") and not only people's perceptions (e.g., implementing NHS CRS felt "like painting the Forth Road Bridge"). Second, she found structural elements that undermined the process: for example, the structured software engineering approach preferred by CFH, which required detailed documentation, clashed with "use and build" flexibility necessary to make the system work (since it was rolled-out before design was complete). Also, she associated a faster turnaround of software improvements with a process of adoption that "seemed more successful" (Vol. 1, p. 98-99). Although the word "seemed" suggests caution on judging "success" (or "failure") and makes a reference to users' own perception of "success", the CR perspective did address directly possibilities for evaluating "what worked" in the given contexts.

The final report is hundreds of pages long, too long to review here in its entirety (Cresswell et al., 2011). As such, we provide only illustrative examples of how the SC and CR perspectives, as interpreted by the researchers, intermingled and complemented each other. They emerge in the writing, but in a subtle way—an unaware reader would possibly not notice. The length of the report was unsuitable for the audience of policy makers, used as they are to short briefs. Therefore, the entire research team had to agree on main points to present in an executive summary, in terms of findings and recommendations for policy. Then, after the report was peer-reviewed, the team was asked to present an even more simplified list of recommendations in the form of a table. We remember now how a SC and CR researcher worked together on the same keyboard and computer screen to make a summary of the executive summary. Care was taken in finding short and direct ways of wording these policy recommendations. The first draft is shown below (Table 4) (the draft was eventually subject to revision by the entire team before publication). Each of the seven recommendations were carefully discussed and agreed by both researchers. Yet, recommendations #1-3 reflect a CR perspective more than a SC one, while the opposite is true for #6-7. For instance, the vision of shared electronic health records was seen as *a* mechanism to support change toward joined-up NHS provision of services (#1); the right software was seen as key for a successful implementation—thus make sure that users are involved in its choice (#2); lack of transparency in contracts and communication among stakeholders was found to be a *cause* of "failure" (#3). A SC perspective is reflected more in #6, where education and meaning come to the fore, matching local and national purposes, and especially in #7, where the concept of working out as an ongoing dynamic process is made explicit. The CR recommendations were therefore made more nuanced and less clear-cut by applying an SC lens.

Table 4. First Draft of Key Recommendations (or Lessons Learned)—Example of Merging SC and CR Insight
Key lessons learned
1. Build on the vision of shared EHR to improve provision of joined-up patient-centered care.
2. Make sure organizations and users are involved in decision making, especially in selection of software systems.
3. Let transparency be key characteristic of a national program, for commercial contracts, stakeholder communication structures, lines of accountability, and responsibility.
4. Support organizations in the implementation and adoption of systems of their choice (financial, material, and expert support).
5. Retain hard won knowledge at both national and local level and make use of skills and expertise in and across sites.
6. Provide possibilities for trusts to educate their implementation teams, users, and other health care organizations as to what EHR means and how it can be strategized to ensure there is a clear local rationale aligned with national purposes.
7. Move away from technology driven implementation models and refocus attention on adoption as an ongoing working-out between staff and technology. Think of technology as an enabler of improved care processes rather than an end in itself.
Source: Initial draft of summary findings prior to final report

4. Discussion

In Sections 3.3.1.and 3.3.2, we use the final report of the evaluation of the NHS CRS as an exemplar case to illustrate how CR and SC assumptions inform analysis and outcomes of evaluation. In this discussion, we focus on two aspects revealed in the presentation of the case above and the literature reviewed: evaluation as judgement and learning, and evaluation as a multi-ontological endeavor.

4.1. Evaluation as Judgment and Learning

Evaluation of IS implementations implies some process of judgement—judgement of "merit and worth", or judgement (and implicitly steps for improvement) of design and implementation (Friedman & Wyatt, 2006, p. 24). In evaluation reports such as ours, judgements are often made in terms of "lessons learned"—that is, lessons to be applied in future actions both in the focal program and beyond. These are extrapolations and generalizations intended to cross contexts and certainly go beyond situated interpretations. Judgement is fundamental to evaluation research but directly raises a potential ontological challenge. This was perhaps the most controversial aspect of our research for CR and SC researchers in our team.

Specifically, SC researchers felt they could not take a position as an arbitrator and judge of the different accounts or understandings of outcomes revealed (Czarniawska, 2001). They treated each interpretation equally, and each could be used individually and in relation one to another to provide different accounts of the NHS CRS. The following statement from one of the papers published from the research summarizes the SC view on the generalizability of insights: "our experience indicated that the implementation and adoption of the NHS CRS could not be sufficiently understood separately and outside of these institutions that shaped and formed it" (Takian et al., 2011). Thus, lessons could not be derived that could claim currency beyond the specific context. The focus of SC evaluation is ongoing and situated, and the job of researchers is to "tell the story" through many voices to offer a rich and insightful account. In this way, a process of learning can be supported as researchers reveal the connection and disconnections between and among accounts and present them back to engaged actors and policy makers in interim reports and presentations. This leaves the final step of learning and suggestions for changing policy or practice to the readers. Thus, the multiple understandings of NHS CRS that we reveal propose a challenge to policy makers and program managers: to reinforce their visions, or to respond to the variety revealed. This is their choice to make. Thus, the outcome of SC evaluation is in presenting participants with a mirror, what Czarniawska (2001, p. 264) describes as a "therapeutic approach"—see also Avgerou (1995). This may be taken further when evaluators and other stakeholders discuss evaluation outputs and jointly consider possible actions for change (Lagsten, 2011). Revealing different interpretations of what constitutes the NHS CRS and different ways in which clinical practice and identity are re-produced and re-constructed through it could establish a strong and distinct basis for explaining how and why different parties responded in different ways to the NHS CRS and, from this, offer a way of understanding the problems experienced during implementation and adoption.

From the CR perspective, in contrast, the primary aim of evaluation is to ascertain "what works under what conditions and for whom" and, in doing so, to contribute directly to the evidence base and to theory building (e.g., a theory of implementation of IS in health care). Such theories are the primary means to capture the mechanisms at work in the real world, and account for the actual (which includes alternative, as yet unrealized and "better" futures).

For CR researchers, the NHS CRS can be seen as an object that exists "out there" in various forms (i.e., as a clinically relevant work practice, as various items of software and as a policy of transformation). In each of these manifestations, the NHS CRS is still seen as open to different interpretations, judgments, and enactments by different stakeholders (e.g., use or non-use of a technology) while retaining its own inherent properties. We had only limited access to the technological artefact (the software packages) and hence limited opportunity to evaluate them "context independently"—for example, according to usability principles or with experimental controls. But we did have "access" to people and their views and experiences through written communications, interviews, and observations. These views and experiences did not "lay somewhere" ready for us to

gather up like products from a supermarket shelf swept into a trolley but had to be sought out, captured, and translated before being analyzed by evaluators. Nevertheless, they were taken as more than "interpretations". They were accepted as facts. Certainly partial and open to contestation, needing triangulation and some context placed around them, but nevertheless facts subject to verification. As facts, and reflecting the realist ontology, they could stand for the empirical reality of NHS CRS. Based on this, implicitly if not explicitly, judgments could be made and lessons proposed— lessons that related to the actual in the Empirical-Actual-Real (EAR) ontology.

In CR, the stratified EAR ontology, and the belief in an (analytical) distinction between agency and structure, provided us with a powerful way of relating what is experienced in the empirical domain (events or outcomes that have been observed) to underlying structures and mechanisms in the real domain—thus helping us to identify structural constraints to NHS CRS implementation and adoption. For example, how the NHS budget arrangements shaped the implementation and adoption of the NHS CRS and led to different local actions.. We could also compare strategies we observed being executed in the world (the empirical) to strategies outlined in policy documents (the actual). Out of these comparisons, judgments could be made concerning what works and what does not work, for whom, and under what circumstances. Stepping back from a "seamless web of practice" provided ontological grounding for addressing the material aspect of the technology (Mutch, 2002; Volkoff et al., 2007) and, for example, enabled us to comment on, and draw conclusions about, the NHS CRS technology design.

These differences between SC and CR accounts, as seen in Section 3, came to be reflected in the final report in the focus and language of different sections. At times we made judgmens and at times we eschewed them. For example, we judged the "fitness" of the technology (in the sections written by CR researchers) but we did not judge the multiple visions of the NHS CRS (presented by SC researchers). But, at the end, for the executive summary, we had to provide recommendations for future action, and, in the end, for better or worse, we all moved towards a CR stance.

4.2. Evaluation as a Multi-Ontological Endeavor

We wrote this paper out of a desire to reflect on our experiences of this evaluation research project as outlined in Section 4.1. The initial proposal for the evaluation did not explicitly emphasize the multi-ontological aspect of the research but, as described, we found ourselves increasingly engaged in some form of multi-ontological endeavor as we moved forward with the analysis and reporting. This prompted some reflections, as reported in this section, on the value of our emerging work as an evaluation and, given we found ourselves so engaged, the principles we could establish under which to undertake such studies.

Multi-ontological studies in many ways are to be deprecated. Evaluation research is hard enough to do without confusion over underlying understandings and assumptions. There should certainly be room in the scientific realm for ontological variety and polite conversation across borders, but not for ad-hoc ontological mash-ups. And yet we see that there are three good reasons why multi-ontological studies are done.

The first is that the world confronts us with complex and distributed socio-technical endeavors that lead us to expect that a multi-disciplinary approach is demanded. And while a multi-disciplinary approach does not necessarily imply a multi-ontological approach, it often will. Thus, in the field of healthcare, the ontology that underpins the dominant discourse of evidence-based medicine may work easily with economics, but it often sits uneasily alongside other social sciences that address institutions, organizations, or cultures. We argue that multidisciplinary research should be explicit and respectful, rather than silent and distrustful (or just oblivious) of ontological difference. Our second argument for multi-ontological studies is that the world changes, any program under evaluation changes, and, as it changes, the research team needs to be open to adopt a different ontological position (which might be more suitable to the changed situation). In our case, delays and increasingly obvious impediments in the implementation of NHS CRS demanded some response from us as evaluators. The evaluation could not concern itself with capturing outcomes if no outcomes (of

implementations) arose. But we understood that "when nothing is happening, something is happening". Hence, a turn to a more processual approach to which the SC ontology seemed a better fit. And of course, if **both** something **and** nothing is happening, then evaluation research may require two journeys, but the travellers will need to keep in contact. So we see the possibility of sequential ontological moves (ontological shifts but not drifts, as Thompson (2011) argues) and studies that create parallel, but ontologically different, accounts. The third argument, perhaps the hardest to make, is for the superior qualities of results that are achieved through some ontological dialogue. We believe that our outputs were more insightful and perhaps more actionable because of the ontological weave they were based on.

Thus, we argue that the process of dialogue between different ontological perspectives is of value in itself. As in the "trading zones" of physics (Galison, 1997), the process of "trading" enriched our understanding. When all researchers come from the same background and subscribe to the same theoretical perspective, their views might be left unchallenged and hence with time become reified. Sharing (trading) accounts and interpretations asked us to reconsider and defend our assumptions and thus, at times, to find ways of accommodating and learning from different views. Occasionally this would result in somebody changing their mind and experiencing an ontological shift (Takian et al., 2011).

But, more importantly in terms of evaluation outputs, as Thompson (2011) suggests, the act of shifting between different ontological positions can generate new theoretical insights. In CR, a belief in an analytical distinction between structure and agency (constrained and enabled by structures) opens a way for action to change existing conditions. In this, the "critical" aspect of critical realism comes to the fore. SC researchers trading with CR researchers may be inspired to develop their critical agenda and be challenged to consider how they can accommodate "macro" concerns (but without resorting to referring to abstract structures, such as market forces, class, or, in our case, medical professions). This can be achieved through detailed accounts of different phenomena that do not aspire to produce grand theories of society, but that provide informative accounts about macro-order phenomena. Although a micro-focus is not an ontological requirement of SC, in practice, due to time and budget constraints, many studies are conducted on a micro-level. In turn, we saw how SC traded focus on making the logic of practice explicit, revealing the legitimizing practices (Czarniawska, 2008) and questioning taken-for-granted assumptions (e.g., of organizational properties, technology characteristics, stability of structures, categories and boundaries). This trade can help to sharpen the critical elements in CR. Hence, we echo Reed's assertion reported in Newton et al. (2011, p. 15):

> [C]*ritical realism's focus on the external/direct structural constraints that are necessarily entailed in the wider domination relations within which governance regimes are institutionally embedded can be complemented by social constructionism's concern with the internalized/indirect normative constraints embodied in the discursive technologies through which subjectivities and identities are fashioned and refashioned.*

Finally, we want to make it clear that we do not postulate that SC can be somehow incorporated in CR. Simply recognizing the importance of subjectivities (e.g., actors' meanings and feelings) and trying to somehow accommodate them in research does not "embed" SC ontological perspective in CR. Essentially, the belief that there is one reality (although never fully knowable) rather than many realities (e.g., products of different discourses) are difficult to reconcile, and it probably should not be attempted (Mir & Watson, 2001). Seeing reality as "out there" and fairly stable (as in CR) versus SC's view of realities as being undetermined (i.e., always in flux and under construction) leads to substantial differences in the approach to understanding the world and to conducting evaluation (and more generally research), which we have tried to illustrate in our examples. Hence, we believe that the two perspectives are not interchangeable and are not able to be compounded. Nevertheless, they

are compatible (Mir & Watson, 2001)[13]. By compatible, we mean that they are jointly useful and capable of producing knowledge that is complimentary and coherent.

4.3. Guiding Principles

In Section 4.2, we discuss our arguments for multi-ontological evaluation. In this section, we turn to how it might be achieved.

First, we must recognize that multi-ontological evaluation is not the same as multi-disciplinary or inter-disciplinary evaluation. Multi-ontological evaluation has to accommodate fundamental differences not just in the aspects addressed but in the way the world (reality–realities) is and is understood. The disciplines involved in multi- or inter-disciplinary endeavors may share the same ontological and epistemological assumptions and thus not face such concerns. Equally, researchers belonging to a single discipline often hold different ontological beliefs and thus undertake multi-ontological studies. Interdisciplinary efforts are usually based on a desire to reach integration and convergence of knowledge and (to some degree) methods (Moran, 2010; Repko, 2008). However, similar efforts applied to different ontologies would produce only inconsistencies and confusion.

In multi-ontological evaluation consensus, synthesis and universally agreed and shared meanings are not the primary goals. We suggest that the power of multi-ontological evaluation research is in dialogue between positions. It does not demand synthesis, but mutual respect for other ontological position is needed, underpinned by the understanding that all parties seek knowledge. Navigating differences in understanding the world and in norms, meanings, and interests may be achieved temporally and locally (i.e. as needed) when fruitful. These ideas draw on Czarniawska's (2008, p. 133) notion of dialogical relationship between practitioners and researchers, and Galison's (1997) concept of the trading zone. Galison's analysis is of how physicists from different paradigms work together with engineers on projects. He describes the trading zone as "a social, material, and intellectual mortar binding together the disunified traditions of experimenting, theorizing, and instrument building" (p. 46). He elaborates on how exchanges can be accomplished and explains that:

> [t]*wo groups can agree on rules of exchange even if they ascribe utterly different significance to the objects being exchanged; they may even disagree on the meaning of the exchange process itself. Nonetheless, the trading partners can hammer out a local coordination, despite vast global differences.* (Galison, 1997, p. 783)

We propose that engaging in trading and dialogue should not undermine fundamental consistency of each ontological position. Multi-ontological integrity can and should be preserved through plurivocality of research designs and outputs. Plurivocality in research, as proposed in Czarniawska (2008, p. 134), refers to plurality of voices of those studied and a distinct but not privileged voice for the researcher. We take this term further to mean plurality of researchers/evaluators' voices. It can be reflected in the choice of evaluation methods and criteria, and in the formative or summative feedback (outputs). Plurivocality in outputs can be achieved in at least two ways. First, by producing separate and distinct narratives, which are nevertheless complementary. For example, papers written from different ontological positions but reporting on the same evaluation project purposely published in the same issue of a journal and referencing each other (Barber, Cornford, & Klecun, 2007; Franklin, O'Grady, Donyai, Jacklin, & Barber, 2007). Second, one carefully crafted account may be produced with different positions and related findings intertwined or juxtaposed but with their ontological origins preserved and made explicit.

In support of this, ontological shifts within evaluation should be undertaken consciously and carefully. When doing so, different perspectives' respective strengths and limitations must be understood and the trade-offs acknowledged. A constant awareness of the need to work for epistemic-ontological alignment within constructs should be maintained (Thompson, 2011). Simple re-labelling (of a process

---

[13] For a more comprehensive discussion of differences and convergences between CR and SC (or social constructivism) refer to the following, among many others: (Al-Amoudi & Willmott, 2011; Mingers, 2004a, 2004b; Mir & Watson, 2001; Mutch, 2005; Newton et al., 2011; Willmott, 2005).

as a noun—e.g., from caring for patients to patient-centered care, or of an entity as verb—e.g., from technology to technology-in-practice) won't quite do.

To maintain this dialogue and to navigate differences, coordination practices are needed. Kellog, Orlikowski, and Yates (2006) suggest that the trading zone might be enacted through practices of display (rendering work visible to others), representation (rendering work legible through use of project genres), and assembly (juxtaposing existing work, reusing prior work, and aligning through provisional settlements). Based on our experience, we suggest that, despite the prevalence of digital collaborative technologies, face-to-face workshops have a key role in sharing work, forging tentative understandings and settlements across ontological perspectives. The form of common project genres, such as interview guides and report templates, might need to be agreed on, but their content (e.g., exact questions to be asked) can remain flexible to preserve plurivocality of research.

Finally, outputs underpinned by different ontological perspectives can offer varied ways of understanding the same phenomena (Thompson, 2011). Tensions that exist between such outputs may be beneficial and lead to new insights (for researchers, participants, and sponsors).

From the arguments above, we distil eight principles for conducting multi-ontological evaluation, beyond the specifics of CR / SC collaborations. In our research, we strived to develop and apply these principles while learning the practice of multi-ontological research.

Table 5. Principles of Multi-Ontological Evaluation Research	
Principle of	**Statement**
1. Key premise	Multi-ontological evaluation is not the same as multi-disciplinary or inter-disciplinary evaluation.
2. Respect	Multi-ontological evaluation should be based on mutual respect among ontological positions, underpinned by the understanding that all parties seek knowledge.
3. Dialogue	The power of multi-ontological evaluation research is in dialogue between positions; it does not demand synthesis.
4. Trading	Navigating differences in understanding the world, in norms, meanings, and interests may be achieved temporally and locally (i.e., as needed and when fruitful) as if in a trading exchange. Consensus and universally agreed and shared meanings are not the primary goals.
5. Integrity	Consistency should be maintained within each position. (Multi-)Ontological integrity can and should be preserved through plurivocality of evaluation research designs and outputs.
6. Transition	Ontological shifts should be undertaken consciously and carefully, avoiding ontological drift.
7. Coordination	Coordination practices (e.g., workshops and common project genres) are needed to maintain the dialogue and navigate differences, while maintaining plurivocal integrity.
8. Insight	Outputs underpinned by different ontological perspectives can offer varied ways of understanding the same phenomena. Tensions between such outputs may lead to new insights (e.g., for researchers, participants, and sponsors).

5. Conclusions

By establishing a national program of evaluations with ten separate projects covering different aspects of the NPfIT, policy makers in the Department of Health explicitly recognized the value of

knowledge that could arise from evaluation, although they may at times seem reluctant to learn from such knowledge (Greenhalgh, Russell, Ashcroft, & Parsons, 2011) or be unclear as to the role of evaluation in the policy process (Cameron et al., 2011). Our evaluation of NHS CRS contributes to the evidence base and, we hope, has informed and will inform policy decisions in England post-NPfIT era and will offer findings and insights with relevance beyond the English context.

In this paper, we show, with examples from our study, how individual researchers' assumptions based either on CR or SC perspectives led them to different ways of conducting and reporting the evaluation. We have illustrated how analysis deriving from one perspective or the other can be applied to the same context (or body of text), producing different but compatible forms of knowledge. We also show that both CR and SC offer insights that can enrich the design, process, and outcomes of evaluation of IS in organizations and more generally any study of IS. Hence, we argue that SC and CR each sensitize evaluators to different aspects under study and invite different evaluation questions and answers. The output of evaluation based on SC and one based on CR are significantly different in their use of knowledge gained in the field, and in the kind of evidence and recommendations that they can provide for policy makers. Despite their differences, which cannot and should not be denied, these two philosophical positions can, when taken together, produce complementary and jointly useful knowledge.

We see that, of the two, the CR philosophy provides foundations on which judgements and abstractions of the form that are expected by policy makers and managers who commission evaluations can be offered. Reflecting on our own experience, we come to the conclusion that the nature of evaluation of large scale IS implementations such as those of the NPfIT, renders a CR approach not only meaningful but also, in a way, inevitable. This is unsurprising since formative and summative evaluations are conducted with a view to making a judgment as to whether some IS-led program of change is working, and in order to improve it or those that follow. A CR evaluation is not intended to produce a clear-cut "pass-fail verdict on a program" (Pawson & Tilley, 2004, p. 365) but is aimed at proposing actions (mechanisms) for improving or intervening in a given situation (often an objective of critical studies—e.g., see Shaw & Stahl (2011)), and to make a case for future action.

In our case, we perceive multi-ontological research to have been beneficial to the evaluation task and outputs, and we argue that multi-ontological evaluation should be welcomed. It can produce nuanced and critical insights, support judgements and recommendations, and stimulate processes of learning that allow insights gained to have relevance within and beyond any specific program under evaluation. Although pluralism in research has been advocated as a way of providing more holistic analysis of a research problem (Davison & Martinsons, 2011), we show in this paper that this general approach requires more than just learning and applying different methods or putting together a multi-disciplinary team capable of researching different aspects. Drawing on Czarniawska's (2008) notion of dialogical relationship and Galison's (1997) concept of the trading zone, we have developed principles for conducting multi-ontological evaluation, where we highlighted the need for mutual respect, dialogue, negotiation, and reflection. These principles and our discussion of CR and SC as a philosophical basis for evaluation form this paper's contributions to literatures on the nature and practices of evaluation and on multi-ontological research. Our story and arguments presented in this paper will, we hope, inform those who consider embarking on similar undertakings, be it an evaluation or any other research endeavor.

Acknowledgements

Ethics and research governance

An NHS Ethics Committee classified this project as a service evaluation (ref. 08/H0703/112). Informed consent from participating organizations and individuals was obtained. We also complied with local research governance requirements. All data has been anonymized and care has been taken to ensure that any quotes in this paper are not attributable to participating individuals.

References

AGoC. (2009). *Report of the Auditor General of Canada to the House of Commons, Chapter 4: Electronic health records*. Office of the Auditor General of Canada.

AHRQ. (2010). *Health IT portfolio annual report.* Retrieved from http://healthit.ahrq.gov/portal/server.pt/community/ahrq-funded_projects/654/health_it_portfolio_annual_report/16758

Al-Amoudi, I., & Willmott, H. (2011). Where constructionism and critical realism converge: Interrogating the domain of epistemological relativism. *Organization Studies, 32*(1), 27-46.

Alvesson, M., & Skoldberg, K. (2000). *Reflexive methodology: New vistas for qualitative research*. London: Sage Publications.

Ammenwerth, E., Graber, S., Herrmann, G., Burkle, T., & Konig, J. (2003). evaluation of health information systems—problems and challenges. *International Journal of Medical Informatics, 71*, 125-135.

Archer, M. (1998). Introduction: Realism in the social sciences. In M. Archer, R. Bhaskar, A. Collier, T. Lawson, & A. Norrie (Eds.), *Critical realism: Essential reading*. London: Routledge.

Ash, J. S., Sittig, D. F., Poon, E. G., Guappone, K., Campbell, E., & Dykstra, R. H. (2007). The extent and importance of unintended consequences related to computerized provider order entry. *Journal of the American Medical Informatics Association, 14*(4), 415–423.

Ashmore, M. (1989). *The reflexive thesis.* Chicago: University of Chicago Press.

Avgerou, C. (1995). Evaluating information systems by consultation and negotiation. *International Journal of Information Management, 15*(6), 427-436.

Barad, K. (2003). Posthumanist performativity: Toward an understanding of how matter comes to matter. *Journal of Women in Culture and Society, 28*(3), 801-831.

Barber, N., Cornford, T., & Klecun, E. (2007). Qualitative evaluation of an electronic prescribing and administration system. *Quality and Safety in HealthCare, 16*(4), 271-278.

Barrett, M., Oborn, E., Orlikowski, W. J., & Yates, J. (2012). Reconfiguring boundary relations: Robotic innovations in pharmacy work. *Organization Science, 23*(5), 1448-1466.

Becker, J., & Niehaves, B. (2007). Epistemological perspectives on IS research: A framework for analysing and systematizing epistemological assumptions. *Information Systems Journal, 17*(2), 197-214.

Bhaskar, R. (1978). *A realistic theory of science*. Sussex: Harvester Press.

Bloomfield, B. P., & Vurdubakis, T. (1994). Boundary disputes: Negotiating the boundary between the technical and the social in the development of IT systems. *Information Technology and People, 7*(1), 9-24.

Blumenthal, D. (2009). Stimulating the adoption of health information technology. *New England Journal of Medicine, 360*, 1477-1479.

Blumenthal, D. (2011a). Wiring the health system—origins and provisions of a new federal program. *The New England Journal of Medicine, 365*(24), 2323-2329.

Blumenthal, D. (2011b). Implementation of the federal health information technology initiative. *The New England Journal of Medicine, 365*(25), 2426-2431.

Blumenthal, D., & Tavenner, M. (2010). The "meaningful use" regulation for electronic health records. *New England Journal of Medicine, 363*, 501-504.

Brennan, S. (2005). *The NHS IT project: The biggest computer programme in the world—ever!* Abingdon: Radcliffe.

Bygstad, B. (2010). Generative mechanisms for innovation in information infrastructures. *Information and Organization, 20*, 156-168.

Cameron, A., Salisbury, C., Lart, R., Stewart, K., Peckham, S., Calnan, M., Purdy, S., & Thorp, H. (2011). Policy makers' perceptions on the use of evidence from evaluations. *The Policy Press, 7*(4), 429-427.

Carlsson, S. (2011). Critical realist information systems research in action. In M. Chiasson, O. Henfridsson, H. Karsten, & J. DeGross (Eds.), *Researching the future in information systems* (Vol. 356, pp. 269-284). Boston: Springer

Carlsson, S. A. (2003). Advancing information systems evaluation (research): A critical realist approach. *Electronic Journal of Information Systems Evaluation, 6*(2), 11-20.

Chen, W., & Hirschheim, R. (2004). A paradigmatic and methodological examination of information systems research from 1991 to 2001. *Information Systems Journal, 14*(3), 197-235.

Cornford, T., Doukidis, G. I., & Forster, D. (1994). Experience with a structure, process and outcome framework for evaluating an information system. *Omega, International Journal of Management Science, 22*(5), 491-504.

Cresswell, K., Ali, M., Avery, A., Barber, N., Cornford, T., Crowe, S., Fernando, B., Jacklin, A., Jani, Y., Klecun, E., Lichtner, V., Marsden, K., Morrison, Z., Paton, J., Petrakaki, D., Prescott, R., Quinn, C., Robertson, A., Takian, A., Voutsina, K., Waring, J., Sheikh, A. (2011). *The long and winding road. An independent evaluation of the implementation and adoption of the National Health Service Care Records Service (NHS CRS) in secondary care in England.* Retrieved from http://www.cphs.mvm.ed.ac.uk/grantdocs/526%20-%20Final%20report%20v31st%20Mar%20FINAL.pdf

Crotty, M. (2003). *The foundations of social research: Meaning and perspective in the research process.* London: Sage Publications.

Czarniawska, B. (2001). Is it possible to be a constructionist consultant? *Management Learning, 32*(2), 253-266.

Czarniawska, B. (2008). *A theory of organizing.* Cheltenham: Edward Elgar.

Danermark, B., Ekstrom, M., Jokobsen, L., & Karlsson, J. C. (2003). *Explaining society: Critical realism in the social sciences.* London: Routledge.

Davison, R. M., & Martinsons, M. G. (2011). Methodological practice and policy for organisationally and socially relevant IS research: An inclusive-exclusive perspective. *Journal of Information Technology, 26*, 288-293.

Dobson, P., Myles, J., & Jackson, P. (2007). Making the case for critical realism: examining the implementation of automated performance management systems. *Information Resources Management Journal, 20*(2), 140-152.

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). New public management is dead—long live digital-era governance. *Journal of Public Administration Research and Theory, 16*(3), 467-494.

Elder-Vass, D. (2008). Searching for realism, structure and agency in actor network theory. *The British Journal of Sociology, 59*(3), 455-473.

Fairclough, N. (2005). Discourse analysis in organization studies: The case for critical realism. *Organization Studies, 26*(6), 915-939.

Farbey, B., Land, F., & Targett, D. (1999). The moving staircase: Problems of appraisal and evaluation in a turbulent environment. *Information Technology & People, 12*(3), 238-252.

Fountain, J. E. (2001). *Building the virtual state: Information technology and institutions.* Washington, DC: Brookings Institution Press.

Franklin, B. D., O'Grady, K., Donyai, P., Jacklin, A., & Barber, N. (2007). The impact of a closed-loop electronic prescribing and administration system on prescribing errors, administration errors and staff time: A before and after study. *Quality and Safety in Health Care, 16*(4), 279-284.

Friedman, C. P., & Wyatt, J. C. (2006). *Evaluation methods in biomedical informatics* (2nd ed.). New York: Springer.

Galison, P. (1997). *Image and logic: A material culture of microphysics.* Chicago: The University of Chicago Press.

Gergen, K. J. (1994). *Realities and relationships: Soundings in social constructionism.* Cambridge, MA: Harvard University Press.

Greenhalgh, T., Humphrey, C., Hughes, J., MacFarlane, F., Butler, C., & Pawson, R. (2009a). How do you modernize a health service? A realist evaluation of whole-scale transformation in London. *The Milbank Quarterly, 87*(2), 391-416.

Greenhalgh, T., Potts, H. W. W., Wong, G., Bark, P., & Swinglehurst, D. (2009b). Tensions and paradoxes in electronic patient record research: A systematic literature review using the meta-narrative method. *The Milbank Quarterly, 87*(4), 729-788.

Greenhalgh, T., & Russell, J. (2010a). Why do evaluations of eHealth programs fail? An alternative set of guiding principles. *PLoS Med, 7*(11), e1000360.

Greenhalgh, T., Russell, J., Ashcroft, R., & Parsons, W. (2011). Why national eHealth programs need dead philosophers: Wittgensteinian reflections on the reluctance of policymakers to learn from history. *The Milbank Quarterly, 89*(4), 533-563.

Greenhalgh, T., Stramer, K., Bratan, T., Byrne, E., Mohammad, Y., & Russell, J. (2008). Introduction of shared electronic records: Multi-site case study using diffusion of innovation theory. *British Medical Journal, 337*, a1786.

Greenhalgh, T., Stramer, K., Bratan, T., Byrne, E., Russell, J., & Potts, H. (2010b). Adoption and non-adoption of a shared electronic summary record in England: A mixed-method case study. *British Medical Journal, 340*, c3111.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). London: Sage Publications.

Hartwig, M. (Ed.). (2007). *Dictionary of critical realism*. London: Routledge.

Hirschheim, R., & Smithson, S. (1988). A critical analysis of information systems evaluation. In N. Bjorn-Andersen & G. Davis (Eds.), *Information systems assessment* (pp. 17-37). Amsterdam: North-Holland.

HoC. (2007). The electronic patient record: Sixth report of session 2006–07. *House of Commons / Health Committee*. London: The Stationery Office.

Horrocks, I. (2009). Applying the morphogenetic approach: Outcomes and issues from a case study of information systems development and organisational change in British local government. *Journal of Critical Realism, 8*(1), 35-62.

Iedema, R. (2007). On the multi-modality, materially and contingency of organization discourse. *Organization Studies, 28*(6), 931-946.

Jones, M. R. (2003). Computers can land people on Mars, why can't they get them to work in a hospital? Implementation of an electronic patient record system in a UK hospital. *Methods of Information in Medicine, 42*(4), 410-415.

Keller, C., Gäre, K., Edenius, M., & Lindblad, S. (2009). *Designing for complex innovations in health care: Design theory and realist evaluation combined*. Paper presented at the Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology, Philadelphia, Pennsylvania.

Kellog, K. C., Orlikowski, W. J., & Yates, J. (2006). Life in the trading zone: Structuring coordination across boundaries in postbureaucratic organizations. *Organization Science, 17*(1), 22-44.

Klecun, E., & Cornford, T. (2005). A critical approach to evaluation. *European Journal of Information Systems, 14*(3), 229-243.

Lagsten, J. (2011). Evaluating information systems according to stakeholders: A pragmatic perspective and method. *Electronic Journal of Information Systems Evaluation 14*(1), 73-88.

Lang, A., & Mertes, A. (2011). e-Health policy and deployment activities in Europe. *Telemedicine and E-Health, 17*(4), 262-268.

Latour, B. (1999). *Pandora's hope: Essays on the reality of science studies*. Cambridge, MA: Harvard University Press.

Lee, A. S., & Baskerville, R. L. (2003). Generalizing generealizability in information systems research. *Information Systems Research, 14*(3), 221-243.

Léger, M. (2000). The Canada health infoway: A vital link to the future. *Canadian Government Executive, 6*(3), 20-22.

Mingers, J. (2001). Combining IS research methods: Towards a pluralist methodology. *Information Systems Research, 12*(3), 240-259.

Mingers, J. (2004a). Real-izing information systems: Critical realism as an underpinning philosophy for information systems. *Information and Organization 14*, 87–103.

Mingers, J. (2004b). Critical realism and information systems: Brief responses to Monod and Klein. *Information and Organization, 14*, 145–153.

Mir, R., & Watson, A. (2001). Critical realism and constructivism in strategy research: Toward a synthesis. *Strategic Management Journal, 22*(12), 1169-1173.

Moran, J. (2010). *Interdisciplinarity—the new critical idiom* (2nd ed.). London: Routledge.

Mutch, A. (2002). Actors and networks or agents and structures: Towards a realist view of information systems. *Organization, 9*(3), 477-496.

Mutch, A. (2005). Discussion of Willmott. *Organization, 12*(5), 781-786.

Mutch, A. (2010). Technology, organization, and structure—a morphogenetic approach. *Organization Science, 21*(2), 507-520.

NAO. (2011). *The National Programme for IT in the NHS: An update on the delivery of detailed care records systems. Report by the Comptroller and Auditor General.* London: The Stationery Office. Retrieved from http://www.nao.org.uk/publications/1012/npfit.aspx

Newton, T., Deetz, S., & Reed, M. (2011). Responses to social constructionism and critical realism in organization studies. *Organization Studies, 32*(1), 7-26.

Nightingale, D. J., & Cromby, J. (2002). Social constructionism as ontology: Exposition and example. *Theory & Psychology, 12*, 701-713.

NORC. (2009). *Health information exchange economic sustainability panel: Final report.* Retrieved from http://healthit.hhs.gov/portal/server.pt/gateway/PTARGS_0_10731_865324_0_0_18/ HIE%20Sustainability%20Paper%20-%20Final%20-2.pdf

Oborn, E., Barrett, M., & Davidson, E. (2011). Unity in diversity: Electronic patient record use in multidisciplinary practice. *Information Systems Research, 22*(3), 547 -564.

ONC. (2012). *Update on the adoption of health information technology and related efforts to facilitate the electronic use and exchange of health information: A report to congress.* Retrieved from http://www.healthit.gov/sites/default/files/pdf/january-2012-update-on-hit-adoption-report-to-congress.pdf

Orlikowski, W. J. (2000). Using technology and constituting structures: A practice lens for studying technology in organizations. *Organization Science 11*(4), 404-428.

Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research, 2*(1), 1-28.

Orlikowski, W. J., & Scott, S. V. (2008). Sociomateriality: Challenging the separation of technology, work and organization. *Annals of the Academy of Management, 2*(1), 433-474.

Oroviogoicoechea, C., & Watson, R. (2009). A quantitative analysis of the impact of a computerised information system on nurses' clinical practice using a realistic evaluation framework. *International Journal of Medical Informatics, 78*, 839-849.

Pawson, R. (2002). Evidence-based oolicy: The promise of "realist synthesis". *Evaluation, 8*(3), 340-358.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation.* London: Sage.

Pawson, R., & Tilley, N. (2004). Realist evaluation. In S. Mathison (Ed.), *Encyclopaedia of evaluation* (pp. 362-367). Newbury Park: Sage.

Payton, F. C., Pare, G., Le Rouge, C. M., & Reddy, M. (2011). Health care IT: Process, people, patients and interdisciplinary considerations. *Journal of the Association for Information System, 12*(2/3, Article 4).

Porter, S., & O'Halloran, P. (2012). The use and limitation of realistic evaluation as a tool for evidence-based practice: A critical realist perspective. *Nursing Inquiry, 19*(1), 18-28.

Potter, G., & Lopez, J. (2001). After postmodernism: The millennium. In J. Lopez & G. Potter (Eds.), *After postmodernism: An introduction to critical realism* (pp. 3-16). London: Athlone.

Reed, M. (2000). In praise of duality and dualism: Rethinking agency and structure in organizational analysis. In S. Ackroyd & S. Fleetwood (Eds.), *Realist perspectives on management and organizations* (pp. 45–65). London: Routledge.

Repko, A. F. (2008). *Interdisciplinary research: Process and theory.* London: Sage.

Robertson, A., Cresswell, K., Takian, A., Petrakaki, D., Crowe, S., Cornford, T., Barber, N., Avery, A., Fernando, B., Jacklin, A., Prescott, R., Klecun, E., Paton, J., Lichter, V., Quinn, C., Ali, M., Morrison, Z., Jani, Y., Waring, J., Marsden, K., & Sheikh, A. (2010). Implementation and adoption of nationwide electronic health records in secondary care in England: Qualitative analysis of interim results from a prospective national evaluation. *British Medical Journal,* 341.

Sauer, C., & Willcocks, L. P. (2007). Unreasonable expectations: NHS IT, Greek choruses and the games institutions play around mega-programmes. *Journal of Information Technology, 22*(3), 195-201.

Sayer, A. (2000). *Realism and social science.* London: Sage.

Schultze, U. (2000). A confessional account of an ethnography about knowledge work. *MIS Quarterly, 24*(1), 3-41.

Scott, P., Prytherch, D., & Briggs, J. (2010). Health informatics: Where's the evidence? *UK Faculty of Health Informatics.* Retrieved Oct 21, 2011, from http://eprints.port.ac.uk/2964/1/Where's_the_evidence.v5.0.pdf

Shaw, M. C., & Stahl, B. C. (2011). On quality and communication: The relevance of critical theory to health informatics. *Journal of the Association for Information Systems, 12*(3), 255-273.

Sheikh, A., Cornford, T., Barber, N., Avery, A., Takian, A., Lichtner, V., Petrakaki, D., Crowe, S., Marsden, K. Robertson, A., Morrison, Z., Klecun, E., Prescott, R., Quinn, C., Jani, Y. Ficociello, M., Voutsina, K. Paton, J., Fernando, B., Jacklin, A., Cresswell, K. (2011). Implementation and adoption of nationwide electronic health records in secondary care in England: Final qualitative results from prospective national evaluation in "early adopter" hospitals. *British Medical Journal, 343,* d6054.

Smith, L. M. (2006). Overcoming theory-practice inconsistencies: Critical realism and information systems research. *Information and Organization, 16*(3), 191-211.

Smithson, S., & Hirschheim, R. A. (1998). Analysing information systems evaluation: Another look at an old problem. *European Journal of Information Systems, 7*(3), 155-174.

Smithson, S., & Tsiavos, P. (2004). Re-constructing information systems evaluation. In C. Avgerou, C. Ciborra, & F. Land (Eds.), *The social study of information and communication technology: Innovation, actors, and contexts* (pp. 207-230). Oxford, UK: Oxford University Press.

Strong, D. M., & Volkoff, O. (2010). Understanding organization-enterprise system fit: A path to theorizing the information technology artifact. *MIS Quarterly, 34*(4), 731-756.

Takian, A., Petrakaki, D., Cornford, T., Sheikh, A., & Barber, N. (2011). Building a house on shifting sand: Methodological considerations when evaluating the implementation and adoption of national electronic health record systems. *BMC Health Services Research, 12*(105), 1-11.

Thompson, M. (2011). Ontological shift or ontological drift? Reality claims, epistemological frameworks, and theory generation in organization studies. *Academy of Management Review, 36*(4), 754-773.

Tsoukas, H., & Chia, R. (2002). On organizational becoming: Rethinking organizational change. *Organization Science, 13*(5), 567-582.

University of Birmingham. (2007). nhs connecting for health evaluation programme (advertisement). Retrieved on October 5, 2013, from http://www.birmingham.ac.uk/Documents/college-mds/haps/projects/cfhep/projects/005/NHSCFHEP005Callforproposals.pdf

Volkoff, O., Strong, D. M., & Elmes, M. B. (2007). Technological embeddedness and organizational change. *Organization Science, 18*(5), 832-848.

Weber, M. (1949). *The methodology of the social sciences* (E. A. Shils & H. A. Finch, Trans., 1st ed.). New York: Free Press.

Williams, R., & Pollock, N. (2011). Research commentary—moving beyond the single site implementation study: How (and why) we should study the biography of packaged enterprise solutions. *Information Systems Research, 23*(1), 1-21.

Willmott, H. (2005). Theorizing contemporary control: Some post-structuralist responses to some critical realist questions. *Organization, 12*(5), 747-780.

Wilson, M., & Howcroft, D. (2005). Power, politics and persuasion in IS evaluation: A focus on relevant social groups. *The Journal of Strategic Information Systems, 14*(1), 17-43.

Wynn, J., Donald, E., & Williams, C. K. (2008). *Critical realism-based explanatory case study research in information systems.* Paper presented at the Proceedings of the International Conference of Information Systems.

Wynn, J. D. E., & Williams, C. K. (2012). Principles for conducting critical realist case study research in information systems. *MIS Quarterly, 36*(3), 787-810.

About the Authors

Ela KLECUN is an Assistant Professor in Information Systems at the London School of Economics. Her research revolves around implications of digital innovations for the way healthcare services are organized, delivered, and consumed, in particular in relation to changes in professional work and identity, and notions of patient-centered care. Her other interests include information systems and policy, digital literacy and exclusion, and the evaluation of information systems. Ela has been involved in a number of funded projects, including the evaluation of the National Care Records Service (NCRS) in England.

Valentina LICHTNER is a qualitative researcher in the field of health information systems, with a background in human-computer interaction, software engineering, and economics. Her main research interests are in the areas of information systems and policy, values and design, patient safety, and in the situated embodied mind and its implications for system design. She has investigated systems and practices around the identification of patients and records, electronic prescribing, integrated care pathways, and workflows. She has conducted ethnographic research in healthcare settings, in both primary and secondary care, and with a variety of stakeholders—from clinicians, administrators, and patients, to pharmaceutical companies and software suppliers. She is currently investigating patients' experiences of acute care. She participated in the evaluation of English National Programme for IT implementations at the London School of Economics.

Tony CORNFORD is an Associate Professor in Information Systems at the London School of Economics. His principal research interests are in open source software, open data, and digital innovation in healthcare. His current research is focused on the supply and use of medicines. This work addresses the consequence of digital technologies for the practices of healthcare and the structures in which it is delivered. He has participated in a number of evaluations of major health information technology projects in England including the National Care Records Service (NCRS) and the electronic prescription service.

Dimitra PETRAKAKI is a Lecturer in Business and Innovation in the University of Sussex. Her research interests revolve around the politics of technology when introduced as a change initiative in healthcare, and other public sector settings, and on its implications for professional work and identity. She is also interested in understanding the process and the nature of organizational change initiated by technology. Her research is situated in the interpretive and social constructionist research paradigm and is conceptually inspired by social study of technology, organization studies, and sociology of health and professionals.