Equity-focused developmental evaluation using critical systems thinking

Citation: Reynolds, M. (2014) "Equity-focused developmental evaluation using critical systems thinking." Evaluation: the International Journal of Theory, Research and Practice **20**(1): 75-95.

Equity-focused developmental evaluation using critical systems thinking

Martin Reynolds
The Open University

Abstract

Developmental evaluation can often question the ethical basis of an intervention in terms of whether it's 'doing the right thing' rather than merely 'doing things right'. Drawing on ideas from critical systems thinking and critical systems heuristics, an evaluation framework with a pro-equity focus is suggested. The framework addresses issues of complexity. It invites theories of change associated with philosophical ethics, and provides a means of surfacing and potentially transforming debilitating relations of power in a complex evaluand. A case study of the long-standing Narmada project in India is used to sketch the workings of the framework. The paper describes how the underpinning methodological ideas of critical systems thinking incorporating triple-loop learning can enhance the practice of developmental evaluation.

Keywords: critical systems heuristics, critical systems thinking, developmental evaluation, ethics, theories of change, triple-loop learning.

Introduction

In evaluating ethical questions of equity - access to resources and the distribution of 'goods' and 'harms' - attention is often diverted from related questions of power (relations) and knowledge (claims). Moreover these questions relate to important wider questions regarding (moral and social) legitimacy. Such ethical and political questions are not easy to grasp or work with in terms of an approach to equity-focused evaluations of interventions. Context clearly matters, so theory-based approaches to evaluation including realistic evaluation (Pawson and Tilley, 1997) and theories of change (Connell and Kubisch, 1998) are both relevant. Attention to purpose and objectives in equity evaluation are also important (Oesterle, 2002). Systems thinking provides an approach that deals explicitly with purposes whilst also enabling more holistic evaluation (Hummelbrunner, 2011). Developmental evaluation, which incorporates some aspects of systems thinking including double-loop learning and complex adaptive systems, provides one particular response to equity-focused evaluations (Patton, 2012).

The task of this article is to suggest how pro-equity evaluation might be enhanced by ideas from a tradition of critical systems thinking (CST). By way of illustration, the paper makes reference to a pro-equity evaluation of the Narmada dams project in India. The original evaluation was undertaken as a short desktop study through an invitation from UNICEF to illustrate the value of systems thinking for equity-focussed

evaluations (Reynolds and Williams, 2012). Some brief extracts are taken from the evaluation for illustrative purposes only. Many of the assertions can be empirically questioned and/or verified but, as with the original evaluation, the task here is not to provide a definitive in-depth evaluation of Narmada. The task is to exemplify features of a critical systems approach in relation to some other complementary evaluation approaches using some brief extracts from the desktop evaluation. Narmada was chosen because of its relatively widespread popular familiarity over a long period of time, and its continual relevance with contemporary widespread ethical and political issues of water security and associated development imperatives.

After a short description of the case study, the paper continues in four parts. The first explores methodological challenges associated with *equity* as an ethical focus of evaluation and its expression through theories of change, and *developmental* evaluation as a learning process for potentially dealing with political issues. Three challenges are explored in the second part in relation to the possibilities offered by *systems thinking* and particularly critical systems thinking. A powerful expression of CST is critical systems heuristics (CSH) - a toolbox of penetrating interrelated questions helpful for evaluating complex situations from different stakeholder perspectives. The third part briefly outlines the application of CSH in relation to Narmada as an illustration of critical systems thinking in practice. The fourth part discusses some of the implications of a pro-equity evaluation based on critical systems thinking

Case study: Narmada project

The Narmada project in India was conceived in the 1940s by India's first Prime minister, Jawaharlal Nehru, but it was not until 1979 that the project took form. It is better described as a long-term programme involving many individual projects associated with the construction of dams along the Narmada River which forms the traditional barrier between North and South India.[1] The project involves the construction of 30 large, 135 medium and 3000 small dams to exploit the waters of the river and its tributaries. Of the 30 large dams, Sardar Sarovar is the largest and most controversial. In 1979, the Sardar Sarovar Project was proposed and attracted initial support from international financial institutions including the World Bank. But after much controversy and protest, particularly since the late 1980s, many financial institutions withdrew support. Protest was led by Narmada Bachao Andolan (NBA), a national coalition movement including people affected by the project, environmental and human rights activists, scientists and academics.

The construction of Sardar Sarovar dam itself was stopped in the mid-1990s. However, in October 2000, the Indian Supreme Court gave a go-ahead again for the construction of the dam. Other dams associated with the wider Narmada project have likewise been developing, come under criticism and have been the subject of protest.

Four general dilemmas emerging from the Narmada project can be summarised:

- Water security: improving water quality and access amidst the threat of water-borne diseases from ensuing stagnant reservoir waters
- Energy security: urban and rural economic development alongside extensive displaced populations from rural areas.

- Food security: change in agricultural practices and shift towards large-scale irrigated farming in the context of a demise of small holdings.
- Sustainability: national economic prosperity alongside threats of ecological impacts particularly the loss of biodiversity in previously rich hydrological systems

Equity issues loom large and potential and actual conflicts in an evaluation of such projects are formidable.

Methodological issues of evaluating equity

Equity is an ethic; a normative judgement regarding the distribution of 'goods' (beneficial effects) and 'bads' (harmful effects) associated with an intervention. Pro-equity interventions – whether *directly* through projects, programmes, and/or policies, or *indirectly* through equity-focused evaluations on interventions – seek to redress uneven distribution. More precisely, a pro-equity intervention would seek to redress the imbalance of goods and bads commonly skewed against stakeholders along lines of, for example, socio-economic class, gender, sexuality, age, physical and mental capacities, or geographic location (ranging from disadvantaged regions of the global South to impoverished local ghettos in any country). Such groups are variously referred to as marginalised, disadvantaged, and/or vulnerable. In the UNICEF publication on 'Evaluation for equitable development results' they are generically termed "worst-off groups" (Bamberger and Segone, 2012 p.3).

Intuitively, ethics is about 'being good' or 'doing the right thing'. So we might say that an ethic on equity is about not treating the worst-off groups badly or wrongfully. Joseph Des Jardins uses the terms **normative** and **philosophical** to distinguish between different ethical traditions (Des Jardins, 2001, pp. 18–19):

> To make ethical judgements, give advice, and offer evaluations of what ought or should be is to engage with *normative ethics* […] Normative judgements prescribe behaviour. 'Pesticide use should be reduced.' 'Factories ought not pollute the air and water.' 'Endangered species ought to be protected.' […] Normative disputes can be frustrating when ethical discussions are left at this level, with disagreements and controversies abounding […] *Philosophical ethics* […] is a higher level of generality and abstraction in which we analyze and evaluate normative judgements and their supporting reasons. This is the level of the general concepts, principles, and theories to which we appeal in defending and explaining normative claims.

Normative ethics deal with value judgements. So for Narmada, some examples of normative judgements might be associated with each of the four key issues (Table 1).

Table 1 Narmada project : normative value judgements on construction of dams

Issues of Narmada Project	Enhanced value *arguments for construction*	Diminished value *arguments against construction*
1 Water security	Supply water to 30m people including drinking water facilities	Increase prospect of insect-borne diseases.
2 Energy security	Improve access to electricity in remote villages.	Dispossess large numbers of poor and underprivileged communities of their land as a source of livelihood
3 Food security	Modernise agricutural practices using irrigated farming	Lose skills in more sustainable farming practices
4 Sustainability (ecological)	Establish wildlife sanctuaries protecting rare species (e.g., Sloth Bear, Wild Ass etc)	Diminish biodiversity through monoculture irrigated farming

But what are the deeper theoretical narratives – the theories of change – underpinning these normative judgements? Philosophical ethics deals more with theoretical underpinnings associated with *doing* what's good (consequentialist ethics), *doing* what's right (deontological ethics), and *being* responsible (virtue-based ethics). Similarly, the three ethical traditions can be expressed in relation to an equity-focused evaluation of Narmada:

(i) A *consequentialist ethic* (e.g. emphasising utilization) considers good and bad (harmful) to be drivers of ethical action. It is the consequences of an action that determine a response to the moral dilemma of whether it is right or wrong.

What are the particular issues that need attention and how might they be related with each other? What are the interrelationships and interdependencies amongst securities for water, energy and food and what particular impact do they have on worst-off groups?

(ii) A *deontological ethic* (e.g. emphasising human rights) considers right and wrong to be independent of consequences. It focuses on the moral dilemma of duty – the rightness or wrongness of actions themselves – as opposed to the consequences of those actions.

How might the key issues be attended to and by whom? Is it just 'them' out there or is it also you/ me/ 'us'? Whose perspectives are relevant to these issues and what realistic role might different stakeholders have in making their perspectives count? How for example may the views of vulnerable groups like pastoralist farmers or other less powerful, and often the most worst-off, members of displaced communities such as women, the disabled, and children, be given expression?

(iii) A *virtue-based ethic* (e.g. emphasising social justice) considers character formation to be a determining factor in addition to either calculations of consequence or the rightness or wrongness of the action itself. It focuses on the moral dilemma of character – virtue or vice (being virtuous or vicious).

Why are some issues privileged more than others, and some ways of dealing with them prioritised over others? What opportunities are there for challenging mainstream ways of dealing with harmfulness and wrongdoing? What attributes of expert behaviour and expert-driven solutions to poverty-alleviation prevail?

Table 2 illustrates some consequentialist and deontological ethical aspects of the Narmada project and some particular virtues and vices associated with each of the four key issues.

Table 2 Ethical issues in the Narmada project

Issues of Narmada Project	**Doing what's good (not harmful)** Consequentialist ethic *Measures of success (impacts)*	**Doing what's right (not wrong)** Deontological ethic *Intentions and obligations (rights)*	**Being responsible** Virtue-based ethics ***Virtues/Vices***
1 Water security	Improve quality of water and access to clean water (avoid disease and drought)	Provide universal access to clean water (not reinforcing or developing skewed access)	Justice/ Injustice
2 Energy security	Improve quality of life for citizens (avoid poverty and use of only economic indices)	Provide opportunity for all humans to flourish (not constraining humans from flourishing)	Moderation/ Greed
3 Food security	Improve range of productive capacities for farming (avoid loss of ecologically sustainable farming skills)	Provide expertise to support appropriate practice (not contriving a simplistic solution)	Humility/ Arrogance
4 Sustainability (ecological)	Improve quality of the natural environment (avoid ecological deterioration)	Provide protection against ecological destruction (not ignoring wider obligations to nature)	Compassion/ Recklessness

While equity is essentially a consequentialist ethic - its primary focus being on impact (distribution of goods/bads) - equity is also clearly related to the deontological ethic of human rights, and the virtue-based behavioural ethic of social justice. An equity-focused evaluation requires attention to normative value judgements about what ought to be, and to underpinning assumptions associated with all three ethical theoretical traditions. So how might developmental evaluation provide guidance towards enabling this ethical focus?

Whilst evaluation is conventionally about *applying* value judgements, evaluation might also be considered as a contributor towards *developing* value judgements. From a *developmental* perspective equity is not regarded as some fixed point of nirvana, subject to endless academic discourse on what constitutes it's absolute essence, but might rather be considered a construct or emergent property in the making.[2]

Developmental evaluation was given expression in the 1990s by Michael Patton as an example of utilization-focused evaluations (Patton, 1994; 2010; 2012). The key idea behind developmental evaluation is that in dealing specifically with complex interventions of change and uncertainty (involving unforeseen events and unintended consequences) there is a need for attending to emergent issues.

The focus in developmental evaluation is on how to change systems. Patton here makes particular reference to the importance of double-loop learning:

> "Social innovators and social entrepreneurs, especially those working on issues of human rights and equity, are typically trying to bring about fundamental changes in systems to change the world. To do so, they have to

> understand how the system they want to change is operating and to make the changes that change the system itself, by getting beyond temporary and surface solutions […] Making changes to improve immediate outcomes is single-loop learning; making changes to the system to prevent the problem or embed the solution in a changed system is double-loop learning" (Patton, 2012 p.105-106).

The contrast with single-loop learning is exemplified by Patton with the quick-fix scenario associated with linear problem–identification–correction processes. So in the case of the Narmada project, single-loop learning might be seen with a superficial viewing that long-standing problems of development are first identified with problems of water security, energy security and food security. Building large-scale dams arguably provide some clear immediate solutions to correct each of the three sets of 'problems'.

Whilst not being literally a 'quick-fix' solution there is an associated technical-fix to such thinking that makes such interventions simple and appealing, particularly if the know-how is readily available.

So what would a double-loop learning approach look like? Given the evidence in India that building large dams can exacerbate inequities regarding access to land, water and livelihoods (e.g., Roy, 1999; Shiva, 2002) it would appear that dam construction might be regarded as a classic exemplar of a 'misguided system' of single loop learning; what Russell Ackoff might call 'doing the wrong thing right' (Ackoff and Pourdehnad 2001). A developmental evaluation approach, coupled with ideas from realistic evaluation, would allow for a more forensic examination of the dynamic interrelationships involved with the 'system' of dam construction, to understand the realities of causation.

However, developmental evaluation is part of a wider set of utilization-focused evaluations where there is a risk of not questioning the purposes of an evaluation. The utility is often that mandated by the client alone: "The ideal is to match the type of evaluation to the situation and the needs of the intended users to achieve their intended uses" (Patton, 2012 p.113). But it is often the stated purposes of clients (as against intended beneficiaries) that need questioning, particularly for an equity-focused evaluation of an existing intervention. For example, until the early 1990s World Bank commissioned evaluations of Narmada had been criticised on the basis that the Bank's prime purpose and agenda is one focussed on the dispersal of funds (i.e. as funding agency) as against its wider mandate in, say, the alleviation of poverty (i.e. as an aid agency). In 1993, under pressure from international activist groups who effectively challenged the prevailing evaluation agenda, the Bank refocused it's evaluation on a pro-equity basis, which resulted in the withdrawal of funding for the Sardar Sarovar project. The risk in any evaluation is in allowing those with dominant voices and interests to control the evaluation agenda, thereby sustaining what is already conceived as being 'the right thing'. The challenge here is not only to question the ethical dimension of what's right, but to also appreciate the political dimension in acknowledging that what's considered 'right' depends on the context and relations of power circumscribing rightness.

A familiar example of deeper refocusing can be found in the challenge to traditions of evaluating famine in India. As an expression of long standing single-loop learning, symptomatic problems of famine are identified – typically in terms of food *shortage* - and corrected, in terms of, say, improving charitable supply and distribution of *more* food. The main issue with such evaluations is in not looking at the underlying multiple and deep rooted causes of famine. The Nobel Economist, Amartya Sen, provided a significant pro-equity evaluation of the 1940s Bengal famine in India with his exploration of entitlements (Sen, 1981). Rather than regarding famine as a result of drought causing a shortage of food – a prevailing perspective on causality at the time - Sen signalled deeper multiple causes of famine associated with the political-economy of India. In effect he took on the task of double-loop learning in seeking out why a system of governance with multiple causalities generates famine. In so doing he was able to take a further reflective step back in identifying questions beyond those relating to the availability of resources (the effects of drought and transport networks etc.) towards relations of power around access to resources; signalling the lack of widespread democratic rights which affected peoples' sense of entitlement.

The insight from Sen is that famines simply cannot occur in democracies. It suggests a shift to a third order level of learning in that it takes account of the relations of power (democratic entitlements) that circumscribes the system of agricultural production and consumption in the context of India.

In sum, an equity-focused evaluation needs to address three interrelated challenges. One is to be cognisant of different interdependent factors through eliciting appropriate ethical theories of change relevant to an evaluand situation in a coherent systematic manner. Utilization represents one relevant theoretical domain, but rights theory and virtue-based theory are also relevant. A second challenge is to invite different viewpoints on what is 'right' regarding the intervention being evaluated and to seek out different rationales for 'rightness' based on issues of equity. The third challenge involves reflecting on the relations of power that often circumscribe ethical judgements.

(Critical) systems thinking

Systems practitioners associated with the evaluation community have identified the influence of three attributes of systems thinking as a confluence of three concepts - interrelationships, perspectives and boundaries (Midgley, 2007; Reynolds and Holwell, 2010; Hummulbrunner, 2011; Hummulbrunner and Reynolds, 2013; Williams, 2013). The three attributes might be used respectively for meeting the challenges of equity-focused evaluation described above. Elsewhere these attributes have been presented in terms of a framework:

> "A critical systems framework constitutes three distinct though interrelated (sub)frameworks: firstly, a framework for understanding …complex interrelationships and interdependencies; secondly, a framework for practice … when engaging with different perspectives; and thirdly, a composite framework for responsibility [and reflection]… in dealing ethically [and politically] with inevitable limitations on being holistically 'universe' and pluralistically 'multiverse'." (Reynolds, 2008 p.385)

This framing of CST has been expressed in terms of a learning device; a systems thinking in practice heuristic (Reynolds, 2011). The heuristic can be more simply understood in terms of three purposeful orientations for the use of CST in any intervention:

(i) *Understanding* interrelationships associated with a situation;
(ii) *Engaging* with contrasting perspectives regarding a situation, and
(iii) *Reflecting* on boundaries of such representations and interactions

Relating these activities to the Narmada project reveals how issues of ethics and developmental evaluation might be addressed from a critical systems perspective.

Interrelationships

Ethical issues of doing 'good', doing 'right' and being virtuous, all with respect to worst-off groups associated with the Narmada project (Tables 1 and 2) can be reconfigured in terms of addressing three interrelated systemic stakeholder questions:

(a) What is at stake? (…doing good)
(b) Who are the stakeholders? (…doing right)
(c) What possibilities exist for improving stakeholdings? (…being virtuous)

Questions on what's at stake may focus on the four general issues - water, energy, and food security, and sustainability - and associated consequentialist issues regarding the impact of intervention – what should happen? Questions regarding agency and the key stakeholders, both involved and affected, relate to rights-based deontological issues – who should do what? Related questions regarding stakeholding address wider behavioural changes manifest in either vicious or virtuous pathways. So for example, what fears might there be for perpetuating existing vicious cycles with entrenchment of inequities amongst different stakeholders, particularly with respect to disparate access to land and water resources? Conversely, what opportunities might there be for developing alternative virtuous pathways that may challenge and change conventional ways of thinking about, say, industrialised agricultural production and/or access to natural resources?

Stakes, stakeholders and stakeholdings clearly interrelate. One thing at stake from, say, changes in agricultural practice through the construction of dams, could be traditional rural lifestyles. This will affect different stakeholders in different ways. From a national government bureaucratic stakeholder viewpoint, the stakeholding could be related to the potential (un)economic nature of patchworks of small landholdings, whereas from an existing farmer-landlord stakeholder viewpoint, the stakeholding might be related to the social and community values of communal property resource management.

Perspectives

Any one 'big picture' or systems map or model gained by an evaluator can only represent a partial perspective. Not just partial with respect to being inevitably incomplete, but partial also in being value-driven and therefore inevitably bias. Developmental evaluation appreciates this point by encouraging stakeholder

participation in the evaluation process. From a 'soft' and 'critical' systems standpoint, such perspectives can be actively expressed, analysed and used as a discursive tool through separate systems modelling. Peter Checkland has been particularly influential in developing systems modelling for generating purposeful discussion involving multiple stakeholders (e.g., Checkland and Poulter, 2006). With this in mind he wanted to simplify the process of understanding different perspectives by using a shorthand form of conceptual systems modelling. Rosalind Armson develops Checkland's systems modelling technique further in simplifying systems of perspective using Checkland's three questions – what? how? and why? (Armson, 2011 pp.213-238). Armson recognises the issue of conflicting perspectives based upon different levels of perceiving the situation.

When evaluating dam construction in Narmada, Armson would ask the question "to what problem is this a solution?" she is signalling a disconnect between a 'what' and a 'why'. The 'solution' provides the 'what' at one lower systemic level, but at a higher level the 'what' in the question is actually asking for a 'why'. Making these explicit can literally help to avoid 'talking at cross-purposes'.

We might for example identify 4 different systems perspectives in Narmada according to particular national State interests; each signalling primary and secondary purposes (Table 3).

Table 3 India State perspectives on the Narmada project

Gujaret

'what'	*Primary*: secure irrigation and drinking water *Secondary*: secure hydroelectric power
'how'	Dam construction ; particularly Sardar Sarovar
'why'	Very poor rainfall and need for more industrial development

Madhya Pradesh

'what'	Primary: prevent water loss to neighbouring States Secondary: limit displacement of villages
'how'	Limited and controlled dam construction with attention to appropriate just recompense measures for displacement
'why'	River Narmada runs mostly through MP. 193 villages out of total of 245 would be submerged by Sardar Sarovar alone

Maharasthra

'what'	*Primary*: secure hydroelectric power *Secondary*: limit displacement of villages
'how'	Build higher dam wall at Sardar Sarovar with attention to appropriate just recompense measures for displacement
'why'	Prominent industrial area but needing a check on rural to urban migration

Rajasthan

'what'	*Primary*: secure irrigation supply *Secondary*: none (not directly in Narmada Valley)
'how'	Build higher dam wall at Sardar Sarovar and build canal network
'why'	Prominent agricultural area in South West but with very poor rainfall

The simplified perspectives captured above are all 'ideal' expressions of different systems, each expressing explicit normative value judgements – 'what ought to be' – from particular State perspectives associated with the Narmada Valley. Other normative positions might similarly be expressed from higher system levels; say, a national or even an international/ global perspective.

Whichever system level is in focus, viewpoints might be further disaggregated in accordance with particular stakeholder perspectives including intended clients (farmers, industrialists, urban or rural communities etc.), decision makers (government agaencies, donor agencies etc.), 'experts' (dam constructors, economists, anthropologists, ecologists etc) and those potentially affected negatively (typically, worst-off groups such as Tribal groups or 'oustees' subject to displacement) by the intervention.

Boundaries

Purposeful systems as described above are expressions of perspectives. As such they are variable and, as with all systems, subject to change. The 'what' and 'why' are not fixed attributes - as in mechanical purposive systems - but rather - as with all human purposeful systems – emergent properties. The emergence arises from system boundaries being continually subject to challenge, modification and revision as circumstances change. The process of *developmental* evaluation can and should assist in this process.

Defining boundaries is an essential part of systems thinking. A boundary differentiates between what is "in" and what is "out", what is deemed "relevant" and "irrelevant", what is important and what is unimportant, who "benefits" and who is "disadvantaged". Making boundary judgements is clearly an ethical activity; attune with double-loop learning. It is also a political activity.

Triple-loop learning in the tradition of CST as described by Flood and Romm (1996) captures the political extension of ethical issues. Whereas single-loop learning questions how existing activities can be done better – relating to the normative ethical dimension in asking *how* we should do what we do, double-loop learning goes one step further and questions whether those activities are the right thing to do – relating to the philosophical dimension of asking *what* things are best to do. Triple-loop learning takes a further analytical step and questions how we know what is the right thing to do or why it is that something appears to be the right thing. Is it primarily influenced by both the power of decision makers – those in control of resources (mightiness), and the power of 'expert'/ rational/ academic argument (rightness)?

In Narmada ethical questions might be raised with respect to the influence of large multinational companies involved with agribusiness in forcing decisions around dam construction using their leverage of financial power, even in the face of expert knowledge advising against intervention because of the ecological damage. Alternatively, expertise itself can be regarded as assuming excessive power. There is considerable expertise around dam construction, particularly amongst multinational building contractors, as well as knowledge associated with other dam constructions.

Such expertise can assume a technocentric power base of arrogant 'rightness' overriding the 'mightiness' of, say, ecological interests and vast numbers of people who stand to be adversely affected by dam construction in the Narmada Valley.

This gives rise to questions of politics; an examination of the relationship between power and knowledge, between 'mightiness' and 'rightness'. So evaluators involved with triple-loop learning might gauge whether the 'right thing' is determined more by some source of coercion or authoritative power of government, or indeed the unquestioned authority of commissioners of evaluations (sometimes referred to as 'decisionism' and/or in terms of autocracy) or determined by some power of authoritative knowledge, expertise and/or righteousness (sometimes referred to as technocentrism and/or in terms of an expertocracy).

Platforms for deliberating on ethical issues can be found at all levels of society, from individual conversations to households, from local communities to a wide variety of regional, national and other international forums. Such political space can be of a less formalised type that support, for example, non-violent direct action, or have more formalised manifestations, as with the establishment of mainstream local, national and international government bodies, private sector affiliations and coalitions of NGOs. The Narmada Bachao Andolan (NBA) coalition provides a particularly significant space for alternative expressions of values, perspectives and ethical traditions.

Critical systems heuristics provides a particularly significant set of tools to enhance pro-equity evaluation using CST

Critical systems heuristics and CST

Critical systems heuristics (CSH) developed by Ulrich (1983) with the significant influence of Churchman (1979) represents one of two recognised strands of CST (Ulrich, 2003)[3]. CSH provides a reference system circumscribed by four sources of influence. In simple terms the reference system addresses issues of:

(i) *values and motivations* built into our views of situations and efforts to 'improve' them (who gets what?);
(ii) *power structures* influencing what is considered a 'problem' and what may be done about it (who owns what?);
(iii) *the knowledge basis* defining what counts as relevant information and skills (who does what?); and
(iv) *the moral basis* on which we expect third parties (i.e., people and/or non-human nature not involved yet in some way concerned) to bear the consequences of what we do, or fail to do, about the situations in question (who gets affected by what some people get?).

In CSH, these four dimensions of a complex situation are called *sources of motivation, control, knowledge, and legitimacy*, respectively (see column 'sources of influence' in Table 4). Each of the four sources of influence have three bounded questions regarding who the stakeholders might be, what's at stake, and what might be the particular stakeholding issues or key problems associated with the particular stakeholder group. Thus there are a total of twelve boundary judgements to be made regarding any situation being examined. The complex situation of interest - for

example, an intervention such as the Narmada project – can be effectively translated into a more manageable system of interest. Table 4 outlines the 12 boundary judgements associated with CSH. Some of the judgements originally phrased by Ulrich are more challenging than others to appreciate, hence my inclusion of alternative wordings in parentheses.

Table 4: Boundary judgements as questions relating to CSH
(adapted from Ulrich and Reynolds, 2010)

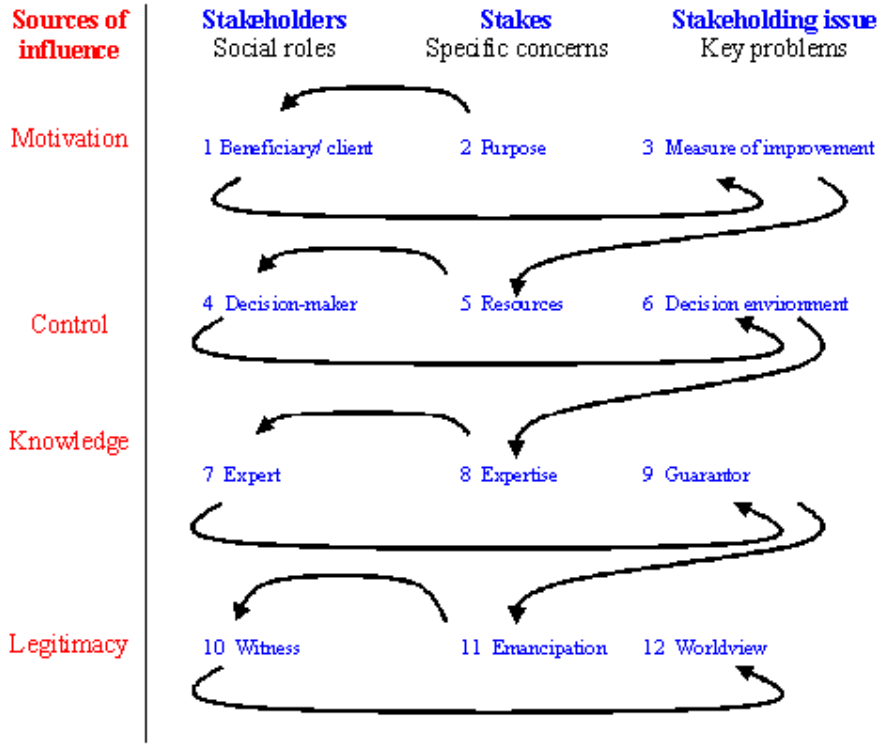
Sources of influence	Boundary judgements informing a system of interest (S) where S may represent an intervention such as a policy, programme or project			
	Stakeholders	**Stakes** *(specific interests)*	**Stakeholding issues** *(key problems)*	
Who gets what? Sources of **motivation**	1. Beneficiaries Intended clients or customers of S?	2. Purpose key objective of S?	3. Measure of success (performance indicators) S's measure of improvement?	
Who owns what? Sources of **control**	4. Decision –makers Those in command of resources necessary to enable S?	5. Resources conditions of success for S - relevant components ('capital') to secure improvement?	6. Decision environment (accountability) conditions of success *outside* the control of the decision maker for S?	'involved'
Who does what? Sources of **knowledge**	7. Experts Those providing relevant knowledge and skills for enabling S?	8. Expertise relevant knowledge and skills supporting S?	9. Guarantor (assurances) promises or guarantee of successful implementation of S?	The
Who gets affected by what some people get? Sources of **legitimacy**	10. Witness ('victims) Those representing the interests of those negatively affected by but not involved with S?	11. Emancipation (marginalisation) constraints on the interests of those negatively affected to have expression and freedom from the worldview of S?	12. Worldview (political space) opportunities available for reconciling contrasting worldviews giving meaning to improvement in S?	'affected'

The following provides a rough sketch on features of CSH as used in the evaluation of Narmada using the parameters of a CST-informed evaluation. The sketch is merely indicative of what a full evaluation might look like. A more comprehensive illustration can be found in Reynolds and Williams (2012).

CSH: understanding interrelationships

The twelve CSH questions prompt an understanding of the 'bigger picture'. In an equity-focused inquiry, it provides a way of organising normative values into a common reference system (sometimes referred to as a 'system of interest'). The relationships between the four sources of influence can be explained through a narrative. Figure 1 below illustrates the narrative in terms of the suggested sequencing between the twelve boundary questions associated with CSH.

Figure 1 unfolding narrative of 12 CSH questions
(adapted from Reynolds, 2007 p. 109)



The narrative in Box 1 below has been adapted for the Narmada project. It was developed originally by Reynolds (2007 p.107) with later modifications from Ulrich (Ulrich and Reynolds, 2010 pp. 260-261) and Williams (Reynolds and Williams, 2012 pp.120-121).

Box 1 Narrative of an unfolding reference system of Narmada associated with CSH (adapted from Reynolds and Williams, 2012 pp.120-121)

Motivation
The design, implementation, and evaluation of an intervention like Narmada starts with some notion of "purpose" e.g., water security. Since a purpose reflects embedded values associated with some person or persons, it is valid to ask, "Whose purpose?" Identifying first what the *purpose* of the system should be helps identify who the intended *beneficiaries* ought to be, e.g. worst off groups, industrialists, farmers, non-human nature…. This in turn raises questions about what should be appropriate *measures of success* in securing some improvement to those beneficiaries. In particular it raises concern about the more 'immeasurable' values such as intrinsic value of nature or historic relational values of community cohesion and traditional practices etc.

Control

The exploration of motivation leads to questions regarding the *necessary resources* or *components needed for success*. Financial capital and other forms of tangible assets like natural, physical, and human capital might be complemented with less tangible factors such as social capital (access to networks of influence). But who might be the *decision makers* in control of such resources? …e.g., government officials, tribal authorities, external funding sources? This in turn prompts questions as to what might be left *outside* the control of such decision makers in order to ensure some level of *accountability*. There are risks of having all resources under the control of the decision makers. External drivers ranging from financial situations, biophysical changes including climate change, are important to Narmada. If the system has all the resources, then the system cannot be held accountable in any way by factors outside the system. What might be part of the decision environment in order to keep the intervention in check and accountable and, importantly, adaptable?

Knowledge

One important set of factors that need to be independent of the decision maker is knowledge or expertise. In an ideal setting, knowledge and expertise ought not to be under the control of the decision maker but rather have independence. So what might be the necessary types and levels of knowledge and experiential know-how to ensure that the system actually has practical applicability and works toward its purpose within the decision making environment? Logistical and economic data is important, but also local rural peoples knowledge and anthropological data might be significant. Who might provide such expertise? The whole point of having experts is to provide some informal warranty or assurance for success. So the question is, how might such expert support provide some promise as an effective guarantor? Conversely this requires evaluators to look out for false guarantors; a reliance on experts or expertise that may turn out to be unwise or misleading. This may range from voluminous so-called unbiased technical reports to provision of nominal participatory exercises such as participatory rural appraisal 'events'.

Legitimacy

Any assessment of the values (motivation), power (control), and expertise (knowledge) associated with any system will always be biased in some way. Expertise can certainly provide some internal legitimacy to the project. But what gives this system the external social legitimacy to carry out its tasks? Churchman (1979) considered that a system could not legitimise itself. Legitimacy is awarded by those outside system. In particular it must withstand *critical* assessment. In other words, if the system is looked at from a different, opposing viewpoint, in what ways might the system's activities be considered as marginalising particular interests? How might it be coercive or malignant rather than emancipatory or benign? Who or what interest groups are likely to be the "victims" of the system – e.g. displaced oustees?... existing resilient ecological dynamics? - and, importantly, what type of representation might be made on their behalf? That is, who might be best capable of making representations on the victims' behalf – tribal representatives, local women activists, international environmental groups…. - and on what basis would they make this claim? Finally, how might the underlying worldview associated with the intervention be reconciled with these opposing worldviews? Where might representation of opposing views be expressed, and what action might happen as a result in order to continually adapt purposes to change in circumstances? The

flourishing of women activist groups in Narmada alongside effective international NGO coalitions have, for example, provided a significant nurturing of political space.

Mapping out the important domains of interrelationships between values, power structures, knowledge claims, and legitimacy, can then invite questions about perspective. Whose views on these interrelationships count?

CSH: engaging with multiple perspectives

CSH boundary judgements can be addressed from different perspectives. For example, each of the State perspectives outlined as simple 'What? How? Why?' systems in Table 3 can each be further flushed out as a full CSH reference system, where the What? …can be aligned with CSH boundary judgements 1 and 2, the How?...can be aligned with CSH boundary judgements 3-9, and the Why? ….can be aligned with CSH boundary judgements 10 to 12.

Each set of CSH boundary judgements can be addressed in a normative 'ought' mode as well as in the descriptive 'is' mode inviting contrasting perspectives through a critique of 'ought' against 'is' – normative against descriptive. Such a critique can itself draw upon different perspectives. Table 5 provides a generic template grid for recording and sketching out such judgements.[4]

Table 5 CSH grid for recording perspectives (adapted from Ulrich, 1996)

CSH Template on perspectives		**Stakeholder** (social role)	**Stake** (role concern)	**Stakeholding issue** (key problem)	
Sources of motivation		1 Beneficiary	2 Purpose	3 Measure of improvement	
	'ought'				
	'is'				
	critique 'is' against 'ought'				
Sources of control		4 Decision-maker	5 Resources	6 Decision environment	
	'ought'				**The involved**
	'is'				
	critique 'is' against 'ought'				
Sources of knowledge		7 Expert	8 Expertise	9 Guarantor/ assurance	
	'ought'				
	'is'				
	critique 'is' against 'ought'				
Sources of legitimacy		10 Witness	11 Emancipation	12 Worldview	
	'ought'				**The affected**
	'is'				
	critique 'is' against 'ought'				

The three traditions of philosophical ethics (Table 2) provide a helpful platform for addressing such critiques in a pro-equity evaluation. For example, CSH questions regarding what's at stake (CSHq2, 5, 8, and 11) can draw on consequentialist ethics regarding the impact of the intervention – the actual effects in prevalence of, say, water-borne diseases, displaced communities, loss of pastoral agricultural skills and knowledge, and diminished biodiversity - from different perspectives – motivation, decision making, expert support, and moral/ social legitimacy. Questions regarding who the stakeholders might actually be (CSHq1, 4, 7, and 10) can draw on deontological ethics regarding who actually have particular rights in the intervention and how such rights/ entitlements might be exercised including, for example, access to water, economic security, livelihood autonomy, and ecological well-being. Questions regarding stakeholding (CSHq3, 6, 9, and 12) can draw on theories of virtue-based ethics addressing whether, and to what degree, particular stakeholders may find themselves entrapped or liberated by their own stakeholder patterns of behaviour, including vicious cycles of injustice (in 'motivation'), greed (in 'control'), arrogance (in 'knowledge') and recklessness (in 'legitimacy').

For an equity-focused evaluation particular attention is given to the perspectives of worst-off groups who traditionally lie outside the core system boundaries (i.e., those affected but not involved – CSHq10-12) in contrast to the perspective of those involved (CSHq1-9).

CSH: reflecting on boundary judgements

Contrasting different stakeholder perspectives through critique can often lead to an unhelpful state of inertia – an entrenchment of stakeholder positioning, or literally 'stakeholding'. CSH boundary judgements 3, 6, 9 and 12 relate to stakeholding development.

Stakeholding development is a positive expression of triple-loop learning. More generally, stakeholding development appreciates the risks in the influence of 'mightiness' (sources of control) over 'rightness' (sources of knowledge) and vice versa. CSH moves developmental evaluation on by recognising different stakeholder concerns regarding possible opportunities for stakeholder development (relating to nurturing purposeful negotiation), as well as signalling risks of stakeholder entrenchment (relating to 'positional bargaining').

The 'stakeholding' issue associated with each of the four sources of influence is a problem of boundary judgement between a bounded system and the realities of the essentially unbounded situation. For sources of motivation the problem is how to make a bounded 'measurement' from the many immeasurable emergent outcomes of an intervention. It questions the politics behind adhering systematically to a reified system of a fixed agenda with fixed targets (and other expressions of performance indicators) rather than allowing for systemic adaptation and revision of purpose and measures in response to feedback during interventions. In Narmada, the recurring 'fixture' appears to be one of dam construction and the whole logic of promoting national economic *growth* in spite of evidence that national growth is not the same as social *development*.

For sources of control the stakeholding problem is how to exert control in an essentially non-controllable socio-economic-ecological environment. It brings to the fore issues of 'might over right' and/or issues of 'might with right'; how much power is unduly expressed by those in control of resources, whether this be international funding agencies or more recently with the growth of Indian economy, national Government?

For sources of knowledge the stakeholding problem is how to give some promise of assurance that the intervention will succeed whilst acknowledging inevitable uncertainty. It brings to the fore issues of 'right over might'; how much power is unduly expressed by experts in the particular field? For example, how much does the multinational interests of the construction industry exert undue influence over governing agencies and bureaucracies in India?

For sources of legitimacy, the stakeholding problem is how to affirm some sense of redress to the power of decision makers and experts about an intervention in a political environment of contested relations of power including contested meanings about righteousness. In Narmada significant political space has emerged in the past twenty years particularly with the emergence of women activist groups, coupled with global activism around dam construction.

Table 6 generically illustrates the contrasting issue of either stakeholding entrenchment or stakeholding development to look out for in addressing boundary judgements in a creative manner.

Table 6 Stakeholding entrenchment or development associated with a system of interest

	Stakeholders *Social Roles*	**Stakes** *Role-specific concerns*	**Stakeholdings** *'Key Problems' = tensions between idealised 'system' vs realities of 'situation'*	
Sources of motivation	1.Beneficiary/ client	2.Purpose	3 – (measure of success) enchantment of fixed *measurable outcomes* vs managing emergence *Check on values (circumscribing the system)*	The involved
Sources of control	4.Decision-maker	5. Resources	6 – (environment) imperative towards *command and control* vs allowing autonomy *Check on power (controlling the system)*	
Sources of knowledge	7. Expert	8.Expertise	9 – (guarantor) dogma and promises of *professional expertise* vs wider humility of social/ ecological uncertainty *Check on complacency (informing the system)*	
Sources of legitimacy	10. Witness	11. Emancipation	12 – (worldview) righteousness and premises of *'the' system* vs rights of, and consequences on, those affected *Check on fundamental meanings (assumed within the system)*	The affected

Discussion

The World Commission on Dams (WCD) was initiated in the late 1990s through the auspices of the IUCN and World Bank. The Commission of 12 appointed experts from different fields was established in response to adverse affects of dam construction, coupled with an extensive period of protest from activist groups, both local and international, in Narmada and other hotspots of similar contentious dam construction programmes such as the Three Gorges in China. WCD reported in 2000 with recommendations for an approach that addresses more explicitly issues of rights and risks (World Commission on Dams, 2000). It remains unclear as to what degree the Bank and indeed other important decision makers have adopted the quite radical recommendations coming from the WCD. How might evaluation practices help to support the adoption of a more rights and risks approach?

Addressing issues of 'rights' in projects like Dam construction invites an ethical pro-equity focus, not just on existing human rights associated with potential displaced oustees from Dam projects, but the rights of future generations and non-human nature. Issues of rights need to be considered in relation to other issues relating to virtue-based ethics and consequentialist ethics (including utilization theory). Addressing issues of 'risks' requires attention to uncertainties; unforeseen events, unexpected consequences. But this should not be done just as a separate constituent study of, say, risk analysis, but in relation to the working influences of underpinning values, power structures involving internal and external drivers, knowledge claims and external legitimacy. Such attention might give better insight to possibilities of systemic success as much as the risks in systemic failure.

Systems thinking is gaining currency in the evaluation field primarily to assess complex interventions. The emphasis has been on understanding how multiple factors and actors within situations behave in relation to each other. Developmental evaluation, for example, embraces a type of systems thinking associated with complex adaptive systems. Such systems are regarded as holistic entities representing 'a bigger picture'. In this paper, coming from a critical systems thinking (CST) tradition, two other attributes of systems thinking are evident. One involves engaging with multiple perspectives. Another requires critically reflecting on judgements made about system boundaries. Such boundary issues relate to potential ethical conflict and associated power relations amongst different entities and/or perspectives.

Questions remain regarding the effectiveness of pro-equity evaluation. Firstly, to what degree is evaluation equipped to actually reveal the interrelationships of complex situations - why interventions work or not, to what effect, for whom, and in what circumstances? Secondly, how might pro-equity evaluation based on a critical systems approach work with other evaluation tools including theories of change, realistic evaluation, programme theory, and values-engaged evaluations? Thirdly, and perhaps most significantly for pro-equity evaluation, to what degree might the practice of evaluation challenge malign relations of power that underpin interventions, and indeed the evaluations of interventions?

Questions of values and fairness are clearly important (Greene, 2001). From a critical systems viewpoint values are closely associated with issue of power, knowledge and legitimacy. Interrelated systemic questions of political economy need addressing. As

with questions of 'value', questions of power, knowledge and legitimacy each need addressing in relation to issues of what's at stake, who the stakeholders might be, and what opportunities and constraints exist for stakeholding development. In Narmada, for example, what relations of power circumscribe agricultural practice? What types of knowledge (economics, anthropology, ecology etc.) are relevant for appreciating the dynamics of development intervention? What forms of legitimacy are conferred through such interventions? In any mapping of interrelationships it is important to take on board and map out composite ethical and political relationships.

So what tools exist in evaluation for generating this wider overview of values in relation to power, knowledge and legitimacy? What's clearly required is a wider and richer picture of the evaluand – a mapping of the situation. In systems thinking the resulting map is itself a system – a human construct – and should not be confused with the actual territory – the situation or evaluand being represented by the system's map or model. Hence 'systems' can be regarded as themselves representative of 'perspectives' on situations. This is a point of departure from viewing systems as real world entities as depicted through everyday language of, say, 'the' legal system, or education system, or finance system etc. It is also a significant point of departure from much of the developmental evaluation literature based on complexity theory (cf. Patton, 2012), as well as wider ideas on 'evaluating the complex' (cf. Forss et al, 2011). The significance lies in the idea that, from a CST viewpoint, all systems are partial. Mapping out interrelationships and modelling perspectives are not neutral activities– someone somewhere decides where to place boundaries, and which of these bounded systems are most important. So any systems design and/or evaluation of a system is partial with respect to being both holistic (what's in and left out in terms of an infinite array of interrelationships in our interconnected world) and pluralistic (whose interests 'count' and whose are discounted in terms of the many perspectives in our multiverse world). Equity-focussed developmental evaluations perhaps need to move on from validity claims based upon being scientifically neutral to being more transparently ethical and political engagements.

Acknowledgement
The article is based on an earlier paper presented at *10th European Evaluation Society Biennial Conference: Evaluation in the networked society: new concepts, new challenges, new solutions* 3-5 October, 2012. Helskinki, Finland.

My thanks to Bob Williams who co-authored the chapter detailing the outputs of a CSH approach to the Narmada Project

Notes

1. [1]Ethical issues of the Narmada case study are discussed more fully in Reynolds (2009)
2. [2] Here I extend the constructivist tradition in 'rights' development as depicted by Michael Freeden (1991), and the developmental idea of the virtue-based ethic of 'justice' as depicted by Amartya Sen (2009).
3. [3] The other strand is called Total Systems Intervention but is not relevant to this paper.

4. [4] The use of such a Table is recommended for drafting individual and/or collective ideas only. There is a problem of using such a template for relaying an evaluation to third parties. Such tick-box exercises diminishes the rich depth of boundary judgements as well as being an inappropriate language tool for meaningful evaluation. It is important to translate the table into a coherent narrative (possibly using a narrative similar to Box 1)

References

Ackoff, R. L. and J. Pourdehnad (2001). "On Misdirected Systems." Systems Research and Behavioural Science **18**: 199-205.

Armson, R. (2011). Growing Wings on the Way: Systems thinking for messy situations. Axminster, UK, Triarchy Press.

Bamberger, M. and M. Segone, Eds. (2012). Evaluation for equitable development results. New York, UNICEF.

Checkland, P. and J. Poulter (2006). Learning For Action: A Short Definitive Account of Soft Systems Methodology, and its use Practitioners, Teachers and Students. Chichester, John Wiley.

Churchman, C. W. (1979). The Systems Approach and its Enemies. New York, Basic Books.

Connell, J. P. and A. C. Kubisch (1998). "Applying a theory of change approach to the evaluation of comprehensive community initiatives: progress, prospects, and problems." New approaches to evaluating community initiatives **2**(15-44).

Des Jardins, J. R. (2001). Environmental Ethics: an introduction to environmental philosophy. Belmont, CA, Wadsworth.

Flood, R. L. and N. Romm (1996). Diversity Management: Triple Loop Learning. Chichester, John Wiley & Sons.

Forss, K., M. Marra, et al., Eds. (2011). Evaluating the Complex: Attribution, Contribution and Beyond. Comparative Policy Evaluation. New Brunswick, Transaction Publishers.

Freeden, M. (1991). Rights. Milton Keynes, Open University Press.

Greene, J. C. (2001). "Dialogue in Evaluation: A Relational Perspective." Evaluation **7**(2): 181-187.

Hummelbrunner, R. (2011). "Systems thinking and evaluation." Evaluation **17**(4): 395-403.

Hummelbrunner, R. and M. Reynolds (2013). "Systems Thinking, Learning and Values in Evaluation." Evaluation Connections: Newsletter of the European Evaluation Society (June 2013).

Midgley, G. (2007). Systems Thinking for Evaluation. Systems Concepts in Evaluation: An Expert Anthology. B. Williams and I. Imam. Point Reyes CA, USA Edgepress**:** 11-34.

Oesterle, A. (2002). "Evaluating Equity in Social Policy: A Framework for Comparative Analysis." Evaluation **8**(1): 46-59.

Patton, M. Q. (1994). "Developmental Evaluation." Evaluation Practice **15**: 311-320.

Patton, M. Q. (2010). Developmental evaluation: Applying complexity concepts to enhance innovation and use. New York, Guilford Press.

Patton, M. Q. (2012). Developmental evaluation for equity-focused evaluations. Evaluation for equitable development results. M. Bamberger and M. Segone. New York, UNICEF**:** 102-114.

Pawson, R. and N. Tilley (1997). Realistic evaluation, SAGE Publications Limited. London

Reynolds, M. (2007). Evaluation based on critical systems heuristics. Systems Concepts in Evaluation: An Expert Anthology. B. Williams and I. Imam. Point Reyes CA, USA EdgePress**:** 101-122.

Reynolds, M. (2008). "Getting a grip: A Critical Systems Framework for Corporate Responsibility." Systems Research and Behavioural Science **25**(3): 383-395.

Reynolds, M. (2009). Environmental Ethics. The Environmental Responsibility Reader. M. Reynolds, C. Blackmore and M. Smith. London, New York and Milton Keynes Zed Books and the Open University**:** 40-51.

Reynolds, M. (2011). Critical Thinking and Systems Thinking. Critical Thinking. C. P. Horvath and J. M. Forte. New York, Nova Science Publishers**:** 37-68.

Reynolds, M. and S. Holwell (2010). Introducing Systems Approaches. Systems Approaches to Managing Change. M. Reynolds and S. Holwell. London, Springer and The Open University**:** 1-23.

Reynolds, M. and B. Williams (2012). Systems thinking and equity-focused evaluations. Evaluation for equitable development results. M. Bamberger and M. Segone. New York, UNICEF http://oro.open.ac.uk/30711/

Roy, A. (1999). The Cost of Living. London, Flamingo (Harper Collins).

Sen, A. (1981). Poverty and Famines:An Essay on Entitlement and Deprivation. Oxford, Clarendon Press.

Sen, A. (2009). The Idea of Justice. Harvard, Harvard University Press.

Shiva, V. (2002). Water wars: Privatization, pollution, and profit, India Research Press.

Ulrich, W. (1983). Critical Heuristics of Social Planning: a new approach to practical philosophy. Stuttgart (Chichester), Haupt (John Wiley - paperback version).

Ulrich, W. (1996). A Primer to Critical Systems Heuristics for Action Researchers. Hull, University of Hull.

Ulrich, W. (2003). "Beyond Methodology Choice: critical systems thinking as critically systemic discourse." Journal of the Operational Research Society **54**(4): 325-342.

Ulrich, W. and M. Reynolds (2010). Critical Systems Heuristics. Systems Approaches to Managing Change. M. Reynolds and S. Holwell (Eds). London, Springer and The Open University**:** 243-292.

Williams, B. and I. Imam, Eds. (2007). Using Systems Concepts in Evaluation: An Expert Anthology. Point Reyes, CA, USA, Edge Press.

Williams, B. and R. Hummelbrunner (2010). Systems Concepts in Action: A Practitioner's Toolkit. Stanford, CA, Stanford University Press.

Williams, B. (2013). "Three core concepts: inter-relationships, perspectives, boundaries." Evaluation Connections: Newsletter of the European Evaluation Society **June 2013**.

World Commission on Dams. (2000). Dams and Development: A New Framework for Decision-making: the Report of the World Commission on Dams, November 2000, Earthscan.