Evaluation as public work: an ethos for professional evaluation praxis

Conference or Workshop Item

How to cite:

For guidance on citations see FAQs.

© 2017 Martin Reynolds

Version: Accepted Manuscript
Evaluation as public work: an ethos for professional evaluation praxis

Authors: Martin Reynolds and Thomas Schwandt

Abstract

The usefulness of evaluation has in recent years been explored through a wider backdrop of conversations regarding the professionalization of evaluation practice. The discourse on professional evaluation has been largely circumscribed around two models of professionalism – social trusteeship and technocratic professionalism. Drawing on the works of Albert Dzur, Thomas Schwandt identifies a third model of democratic professionalism which promises a more purposeful discourse; a means of rendering the usefulness of evaluation as part of a civic professionalism in terms of what Henry Boyte calls ‘public work’. Whereas mainstream democratic professionalism emphasises the need for deliberation, the idea of public work emphasises the democratizing ideal of co-creativity with the active involvement of evaluators and citizens as co-agents of change. According to Schwandt, what is missing is an appropriate framework for justifying democratic professionalism as the ethos for professional evaluation practice; an ethos that relates to elements of ethics and legitimacy. Such a framework, Schwandt claims, might be appropriated from systems thinking and the tool of critical systems heuristics as used and developed in evaluation practice by Martin Reynolds. This paper explores the usefulness of evaluation through systems thinking in practice as a mediating device; a purposeful model that converses between the use of evaluation amongst citizens (including commissioners of evaluations) in an evaluand, and the usability of evaluation amongst evaluators and citizens as co-creating agents. The triangulation can be expressed through the praxis of evaluation; a praxis that at best might constitute a civic professionalism based on an ethos of evaluation as public work.

Key words: civic professionalism, critical systems heuristics, developmental evaluation, modelling, public work, systems thinking in practice

1. Introduction: usefulness of evaluation

At a conference in March, 2017, entitled ‘Sustainability in Turbulent Times’ hosted by the Nexus Network in the UK several recurring themes and questions relevant to the usefulness of evaluation practice emerged. Firstly, why is systemic failure in the policy domain so prevalent and how might practices (including evaluation) be used better to avoid policy failure? A second related theme emerged regarding the institutional failure – usability - of research and policy. A particular recurring
concern on this matter was recognising the value of interdisciplinary and transdisciplinary research and scholarship. For the large part, it was claimed, research excellence and accountability remains fixed in rigid conventional disciplines. What rooms of manoeuvre exist for, say, evaluation practice to thrive as a transdiscipline, and indeed what skills and techniques need to support such transdisciplinarity (cf. Scriven, 2008)? A third recurring theme of relevance to evaluation from the conference is associated with the populist scepticism about expertise arising from and within the 2016 Brexit referendum and Presidential election in USA. Several delegates and speakers raised questions regarding media-hype on the so-called ‘post-truth’ society, preferring instead to signal evidence more of a ‘post-trust’ society. Decision makers and experts supporting decision making - including evaluators - are lacking the societal trust and social legitimacy needed to perform useful practice. Different attributes of ‘trust’ are needed that go beyond some scientific notion of seeking the truth.

The three interrelated themes speak respectively to issues of (i) the evaluand – systemic failures and questions of use of evaluation amongst those subject to, and users of, evaluation; (ii) evaluators with other stakeholders – institutional failures and questions regarding the ‘usability’ of techniques in evaluation; and (iii) evaluations themselves – societal failure and questions regarding the social legitimacy and ‘usefulness’ of evaluation. This paper explores the usefulness of evaluation through the prism of conversations around the professionalism of evaluation.

In a concluding section of a draft paper entitled “ What Is Missing in Efforts to Professionalize Evaluation Practice?” Thomas Schwandt noted the following:

“I believe it is possible to appropriate and justify democratic professionalism as the ethos for professional practice by drawing on the work of Ulrich and Reynolds on critical systems heuristics [Schwandt, 2015] but that is a matter for another paper”... Schwandt (2016 p.21)

Four contributory lenses influencing a model of civic professionalism for evaluation can be identified, associated with the writings of Dzur, Boyte, Schwandt, and Reynolds. Albert W. Dzur (2008) identifies three models of professionalism. The earliest model is social trusteeship, where professionals work on behalf of citizens giving voice to their concerns, though remaining dispassionate in the process. A second model of technocratic professionalism emerged in the 1960s with the growth of professional bodies and distrust of ruling elites. Technocratic professionalism is characterised by Dzur as being disempowering of democratic publics – convincing decision makers that lay expertise is not to be trusted. Technocentrism is also depoliticising, by converting complex political issues into solve-able problems using the skills of the profession, thus making allegiance to government administration and bureaucracy rather than civil society. The third model identified by Dzur is democratic or civic professionalism; a revitalisation of earlier social trusteeship but where beneficiaries are not regarded as clients or customers but rather as citizens, each with a stake in professional decision making. Democratic professionalism “seeks public good with and not merely for the public (ibid, p.129 original emphasis).

Harry C. Boyte (2011; 2014) provides a potential ethos to underpin democratic professionalism in demarcating a difference between deliberative democracy and ‘public work’. Drawing on the work of John Forester, Boyte signals the problems of the deliberative practitioner in focusing excessively on language “We always face the danger that we will listen to what is said and hear words, not power; words, not judgement; words, not inclusion and exclusion; ‘mere words’ and not problem-framing” (Forester, 1999 cited in Boyte, 2014 p.1). Boyte’s notion of public work goes beyond deliberation as a strand of civic practice and invites more a sense of civic agency, where professionals and citizens act together in order to co-create the world rather than just deliberate about the world.
Thomas Schwandt brings together various ideas of evaluation as professional practice (Shwandt, 2014; 2015a; 2015b; 2016; Schwandt et al., 2007; Schwandt and Schwandt, 2015). With particular attention to ideas from Dzur and Boyte, Schwandt draws out an ethos of professional evaluation practice based upon public work (Schwandt, 2016). In a wide ranging synopsis of various endeavours towards the professionalization of evaluation, including reference to several codes of ethics established by evaluation organisations including, amongst others, the UK Evaluation Society (UKES Guidelines for Good Practice in Evaluation), the Australasian Evaluation Society, the African Evaluation Association, the American Evaluation Association, and the Canadian Evaluation Society, Schwandt (ibid) signals a missing underpinning ethos based on an agreed ethic. “Although these organisations endorse various standards for evaluation practice (most modelled on the broad principles of utility, technical competence, cultural competence, proper treatment of individuals, fiscal responsibility etc.) commentary on the civic purpose of evaluation – the manner in which the public interest is to be understood and served by evaluation – is relatively meager” (ibid p.18).

Moreover, there would appear to be considerable difficulty in gaining some consensus of such purpose given the disparate interests amongst evaluators. Drawing on the more generic work of Brint (2015) regarding pitfalls in developing civic professionalism, Schwandt concurs that a civic professionalism of evaluation is undermined by three different ethical traditions variously at play amongst evaluators – an ethos of care and nurturance (serving the underserved), of duty (performing one’s duty in view of some transcendent principle such as respect for evidence, skepticism with respect to truth claims, etc.), and an ethic built around the reciprocal benefits of exchange (if the professional adds value, then clients confer benefits to the professional that provide for stable practices and good salaries)” (Schwandt, 2016 p.21). The best that might be achieved in developing a civic professionalism in evaluation practice based on public work is to have some means of continual conversation around such ethical positions. As noted earlier, Schwandt sees the possibility of such a mediating device drawing on traditions of systems thinking and critical systems thinking.

From a perspective of critical systems thinking (CST) Martin Reynolds has variously contributed towards discourse on expert support (Reynolds, 2008a; 2011), corporate responsibility (Reynolds, 2008b; 2014a) and systems-based evaluation (Reynolds, 2014b; 2015a; 2015b). Core to this work has been development of a framework of systems thinking in practice based on CST for nurturing purposeful action-oriented conversation. CST builds on conventional systems thinking dimensions of (i) understanding inter-relationships and (ii) engaging with multiple perspectives, by way of (iii) reflecting on boundary judgements. Boundaries are what constitutes ‘systems’ – how we bound particular sets of inter-relationships and bound particular perspectives. Reflecting on boundary judgements means dealing critically with the partiality of (a) inevitably not incorporating all inter-relationships, and (b) inevitably being biased in dealing with perspectives. CST thus draws attention to the responsibility of practitioners and hence speaks to the possibilities of nurturing conversation towards civic professionalism in the field of evaluation. One significant set of systems ideas/tools used in CST is critical systems heuristics (CSH), originally developed by Werner Ulrich (1983), and further elaborated upon using case study evaluations by Ulrich and Reynolds (2010). CSH uses a template for exploring the ethics, power, knowledge, and legitimacy of any situation. It is therefore a potentially helpful device for surfacing and deliberating upon the missing elements of an ethos of public work alluded to by Schwandt.

This paper uses ideas of systems thinking for framing the usefulness of evaluation through the prism of civic professionalism. Firstly, after providing an initial short working definition of systems thinking, the three types of professionalism outlined by Dzur are formally ‘modelled’ (framed) using ‘simple’ systems. A second more granular level of modelling civic professionalism is then introduced through CSH. As a ‘reference system’ CSH flushes out some of the missing elements of the civic
professionalism identified by Schwandt based on principles of ‘public work’ expressed by Boyte. A third more generic level of purposeful systems modelling based on ‘systems thinking in practice’ is then discussed from a methodological perspective. The resulting model of evaluation praxis suggested builds on ethical traditions of value judgements underpinning evaluation, and comprises a built-in ethos of public work. The model of praxis provides a contribution towards enabling further development of a model of professionalism for evaluation.

2. Modelling-1: (simple purposive) systems of professional evaluation

Put simply, systems thinking provides a way of rendering real-world complex issues (for example ‘evaluands’ such as policies/programmes/projects etc.) regarded as ‘situations of interest’ into conceptual constructs or ‘systems’ of interest. These rendered systems can then be used to understand relationships better, and to bring multiple perspectives together to design improved systemic interventions that create improved situations.

Any complex situation can be rendered into a system by merely ascribing some purpose. For example, the complexities around professionalism of evaluation might rendered as possibly:

(a system to) ... regulate evaluation practice
... protect careers in evaluation
... remediate problems of incompetence amongst evaluators
... elevate evaluation as recognised academic discipline
... secure assurances amongst commissioners and society at large
... justify large fees for services provided
... nurture a protective safe space for evaluators to develop skills
... set standards of compliance
... effect occupational closure and control over methods and skills ...

Such systems might be called ‘snappy systems’. They can be useful in first delineating (rendering) various (often contrasting) perspectives around a particular situation. Any snappy system can be further elaborated into a ‘simple system’ where the ‘what?’ question regarding purpose might then be unfolded towards a ‘why?’ question regarding an underlying rationale or reasoning or worldview that gives the purpose some legitimacy. Once the what and why are established, a further ‘how?’ question might be asked, elaborating on the processes needed to achieve the ascribed purpose (the ‘what?’) in coherent alignment with its rationale (the ‘why?’). Such simple systems can then be described in terms of simple systems regarding ‘What? How? Why?’ using the generic formula ‘a system to do….what?, by means of….how?, in order to….why?’ (Checkland and Scholes, 1999, Armson, 2011).

The three ‘models’ of professionalism outlined by Dzur (2008) might thus be expressed as rendered simple systems (Figure 1)

Social trusteeship (classic model of professionalism)

| ‘what’ | A system to serve the public (the Other) as ‘clients’ |
| ‘how’ | ...by means of specialized disinterested knowledge and skills |
| ‘why’ | ...in order to protect (but not involve) the public from market interest and political interests |

Technocratic (expert-led model of professionalism)
Democratic (civic-led model of professionalism)

<table>
<thead>
<tr>
<th>‘what’</th>
<th>A system to serve the public as (participant) ‘citizens’ (rather than as residual ‘clients’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘how’</td>
<td>…by means of sharing tasks through meaningful ‘conversation’</td>
</tr>
<tr>
<td>‘why’</td>
<td>…in order to transform, renovate, reconstruct, co-create, empower, civil society</td>
</tr>
</tbody>
</table>

Fig. 1 Three (simple system) models of professionalism (based on interpreted work of Dzur, 2008)

The ‘What? How? Why?’ models in Figure 1 are rendered systems depicting some prevailing perspectives on current discourse relating to the usefulness of evaluation. In each case the ‘usefulness’ of evaluation can be aligned with the higher level purpose, ‘why?’ question. The lower level purpose question regarding ‘what?’ might similarly be aligned with issues regarding the ‘use’ of evaluation, whilst questions of ‘usability’ might be aligned more with the ‘how?’ question. As a departure from ‘snappy systems’, simple systems modelling particularly surface these ‘how’ issues.

For example, in framing evaluation as democratic/civic professionalism, some issues regarding usability might be identified. Firstly, the model clearly privileges some sets of evaluation practices over others. In particular, participatory evaluation is perhaps given prominence through injunctions of ‘sharing tasks’. One of the most prominent techniques associated with democratic professionalism might be deliberative democratic evaluation (House and Howe, 1999). But does that render other techniques and skills of evaluation as being of little value? Are there not circumstances where skills and techniques associated with social trusteeship or technocratic models of professionalism might be legitimate? Schwandt (2016) identifies two other associated issues with the ‘how’ questions of civic professionalism for evaluation; particularly relying on specific evaluation techniques associated with deliberative evaluation. One issue relates to the role of evaluators. Techniques like deliberative evaluation regard evaluators as being custodians of those without a voice or least well-off. In contrast, civic professionalism as described by Schwandt...

“...would not assume an ideological, emancipatory commitment on the part of the evaluator or an a priori advocacy stance in favor of the particular ways of thinking and acting of some specific group of citizens or stakeholders. What it would represent is a way of working with evaluative knowledge in everyday situations of debate and decision making. That approach is emancipatory only in the sense that, and to the degree which, it enables evaluators, stakeholders, and citizens to engage in learning, solution questioning, and successfully arguing their concerns in a variety of local, regional, and national arenas.” (ibid, p.17).

A second issue identified by Schwandt in relation to adopting deliberative evaluation practice reflects the more general concern from Boyte regarding the over-emphasis on ‘deliberating’ rather than ‘producing’. In other words, a civic professionalism requires processes that are governed by purposeful co-creation amongst evaluators and citizens.
Aside from these ‘how’ questions, there are other concerns regarding the general discussion on professionalism in evaluation. Regarding the ‘what’ and ‘why’ questions (of use and usefulness respectively), the professionalism of evaluation practice arguably lacks the conversation involving ethics and legitimacy: “an ethos - the group’s distinguishing moral character and its sense of the civic purpose it serves” (Schwandt, 2016 p.11). Critical systems heuristics (CSH) provides what Ulrich calls a ‘reference system’ (Ulrich and Reynolds, 2010) to render such issues of ethics and politics more explicitly.

As with simple systems, a CSH reference system includes elements of ‘what’ (use) in terms of an immediate purpose, but which can potentially elaborate more on ethical dimensions. CSH can probe more on the ‘values’ associated with any claimed purpose. CSH reference system also elaborates more on the ‘why’ (usefulness) question in terms of flushing out an underlying rationale or worldview, but elaborated in a political context of contrasting worldviews. Moreover, the ‘how’ question in CSH does not preclude relevance of particular techniques, but rather opens them up for critical appraisal. The usability issues are further demarcated in terms of a more differentiated set of questions relating to different stakeholder perspectives.

3. Modelling-2: (complex purposeful) systems for professional evaluation

CSH might be called ‘complex’ as it renders different stakeholder perspectives, including intended beneficiaries, decision makers, experts of different kinds, and victims (those affected negatively by the system but not involved with the system)

The notion of a ‘reference system’ in CSH suggests a sense of built-in provisionality; a system subject to change and development. A reference system speaks to the more fluid change-able notion of ‘purposeful’ systems. Shifting from purposive to purposeful conversation represents one challenge towards expressing civic professionalism of evaluation. A second challenge for enabling purposeful conversation on civic professionalism is in capturing the ethical and political dimensions of professionalism.

CSH renders complex situations into a reference system through asking twelve questions. Figure 2 depicts the twelve CSH questions (hence referred to below, in notation form ‘CSHq1-12’) in a grid with alignment along two parameters: vertically through four sources of influence that affect any situation (rendered as a system) – sources of motivation, control, knowledge, and legitimacy; and horizontally through three sets of stakeholder attributes – relevant stakeholder roles (intended beneficiaries, decision makers, ‘experts’, and witnesses to those affected by the system), relevant stakeholder concerns, and relevant stakeholder problems (stakeholdings associated with the stakeholder role) (Fig.2).
Rendering ideas of professional evaluation through CSH reference systems surface ethical and political issues in several ways. Firstly, all twelve of the CSH questions can be addressed both descriptively (‘is’ mode) and normatively (‘ought’ mode); so a reference system can be used for both empirical analysis and for design. Used in design mode, CSH engages with normative ethics. Secondly, as described elsewhere (Reynolds, 2014b), a deeper ethical endeavour using CSH is also permitted through asking what’s good/ bad (in terms of consequentialist ethics) what’s right/wrong (regarding deontological or rights-based ethics) and what might be behaviourally benign/malign (regarding virtue-based ethics). Such deeper ethical questions are particularly relevant in regard to CSHq 1-3 (sources of motivation) and CSH q10-12 (sources of legitimacy).

As signalled in Fig.2, the four stakeholder groups can be further delineated between those ‘involved with’ a system (beneficiaries, decision makers, and knowledge providers/ experts) and those ‘affected by’ a system. The significance here is an ethical one as well as a political one. Ethically, it builds on West Churchman’s insight that any system will inevitably generate victims (or ‘enemies’ as Churchman called them) – those people and/or associated ideas that are negatively affected by a system (Churchman, 1979). The injunction here is to always look out for victims or enemies of a system (where a system of interest might represent for example an intervention such as policy, programme, or a project etc.) rather than presuming ‘win/win’ scenarios as so often favoured by ‘experts’ (e.g. consultant evaluators) misleadingly offering guarantees (CSHq9) to decision makers (CSHq3 e.g. ‘politicians’ or commissioners). It is ethically misleading to presume that only ‘good’, ‘rightfulness’, or ‘virtue’ will arise from a system of interest irrespective of the perspective taken.

The involved vs affected also raises an important political issue. The term ‘politics’ is used here in the lower-case ‘p’ sense regarding power relations, and the general playing out of differences in power (as against the more specific upper-case ‘P’ of political party politics). Drawing on CSH, relations of power are manifest in different ways. In addition to general tensions between the involved and the affected, tensions exist for each of the stakeholding questions (CSHq3, 6, 9 and 12). Measures of success (CSHq3) for example signal tensions between measurable and non-measurable...
outputs/outcomes/impacts. The decision environment (CSHq6) signal tensions between control over components and adaptation to externalities. Guarantors (CSHq9) signal tensions between an emphasis on technical certainty/reliability/truth, and less certain assurances relating to relevance/pragmatism/trust. Worldviews (CSHq12) signal tensions between the rationale underpinning a system of interest and a contrasting rationale associated with those affected negatively by a system. CSH reference systems can also be used politically in contrasting the two judgements – normative value judgements of ‘ought’ with empirical ‘is’ judgements of ‘fact’. In other words the CSH reference system can be used ‘politically’ by enabling critique.

This paper does not have space for a full-blooded rendering of a normative system of evaluation based on civic professionalism. Figure 3 hints at some of the possible conversations to be had in such a rendition; focusing on the stakeholding categories (CSHq 3, 6, 9, and 12) of a CSH reference system.

<table>
<thead>
<tr>
<th>Sources of influence</th>
<th>Boundary judgements informing professional evaluation practice (CSH-lite overview focusing on stakeholding issues)</th>
<th>Stakeholding issues being discussed (key problems) ...’is’ mode</th>
<th>Stakeholding issues that could be discussed (key problems) ...’ought’ mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developing value</td>
<td>CSHq3 measures of success&lt;br&gt;Democratic model...</td>
<td>(ethical space of purposefulness) from single loop learning (are we doing things right?), to double loop learning (are we doing the right thing?) and triple loop learning (is rightness controlled by might and righteousness controlled by righteousness) ... discussion on ethical issues – social purpose of professional evaluation practice?</td>
<td>(control of power-over) +... power-with (other stakeholders); power-to (effect change); and power-within (empowerment)</td>
</tr>
<tr>
<td>Sources of motivation</td>
<td>Developing power&lt;br&gt;CZHq6 decision making environment&lt;br&gt;Social/trusteeship model</td>
<td>Regulation of evaluation practice; standardisation; managerialism; decentralised control; definitional concerns; Conversation dominated by ‘power-over’ discourse.</td>
<td>(guarantors of reliability) + ... ‘rigour’ including complementary co-guarantors of resonance (response-abilities), and relevance (to social-ecological improvement)</td>
</tr>
<tr>
<td>Developing know-how</td>
<td>CSHq9 guarantors/assurances&lt;br&gt;Technocratic model</td>
<td>Content of evaluation: certifying; credentialising; licencing; competence framework; cross-disciplinary and transdisciplinary nature of evaluation; logic model of evaluation (criteria &gt; standards &gt; measure &gt; synthesise) etc.</td>
<td>(guarantors of reliability) + ... ‘rigour’ including complementary co-guarantors of resonance (response-abilities), and relevance (to social-ecological improvement)</td>
</tr>
<tr>
<td>Sources of knowledge</td>
<td>Developing governance&lt;br&gt;Worldviews/ rationale</td>
<td>(political space of deliberative democracy) opportunities available for reconciling contrasting worldviews regarding professional evaluation practice; incorporating principles of humility, empathy, fallibility, circumscribed by civic professionalism virtue of social-ecological justice</td>
<td>(guarantors of reliability) + ... ‘rigour’ including complementary co-guarantors of resonance (response-abilities), and relevance (to social-ecological improvement)</td>
</tr>
<tr>
<td>Sources of legitimacy</td>
<td>Democratic model...</td>
<td>(ethical space of purposefulness) from single loop learning (are we doing things right?), to double loop learning (are we doing the right thing?) and triple loop learning (is rightness controlled by might and righteousness controlled by righteousness) ... discussion on ethical issues – social purpose of professional evaluation practice?</td>
<td>(control of power-over) +... power-with (other stakeholders); power-to (effect change); and power-within (empowerment)</td>
</tr>
</tbody>
</table>

**Fig. 3. What’s missing in conversations on professional evaluation practice**

Some hints towards conversations that can be had using critical systems heuristics (CSH) questions regarding stakeholding issues associated with sources of motivation (CSHq3), control (CSHq6), knowledge (CSHq9), and legitimacy (CSHq12) © Martin Reynolds, 2017

Figure 3 includes suggestions of what is currently being discussed, particularly using prevailing models of professionalism, identified here in relation to sources of control (social trusteeship) and sources of knowledge (technocentrism). Fig. 3 also prompts suggestions on what ‘ought’ to be discussed more in relation to models of civic professionalism in relation to stakeholding issues associated with all four sources of influence. This partial rendered CSH system attempts to signal what Schwandt (2016) earlier identified as what’s missing in such conversations.
So how might conversations on professionalism in evaluation continue? What further use might be gained from systems thinking/ modelling? Whilst ‘simple systems’ and ‘complex systems’ modelling are helpful in rendering perspectives/discourses as models, the systems generated are at risk of remaining quite static (i.e. remaining ‘purposive’ as against ‘purposeful’). The reality of any discourse is that they remain in a constant state of flux. Schwandt (2016) makes the point that in the sphere of evaluation practice, the three discourses are interlinked and variable rather than existing as discrete entities. The CSH reference system can capture such linkages between contrasting discourses through identifying them as constituent sources of influence on a system (as depicted in Fig. 3).

Another level of systems modelling captures an ongoing conversation in terms of systems thinking in practice (STiP). The model of STiP is deeply purposeful – comprising of dynamic processes where the purposes may change in adaptation to changing circumstances. Here the challenge regarding usefulness of evaluation might be seen in relation to understanding evaluation as praxis (theory-informed-practice, and practice-informed-theory).

4. Modelling-3: evaluation as praxis for public work

“The public work framework reimagines settings like schools, colleges, businesses, government agencies and others not as part of a static “system world” but rather as human inventions, communities which can be reorganized to become enabling environments for agency. This requires conceiving work itself as activity whose public dimensions can be deepened. [...] A deliberative public work approach to democratic change differs from a service framework, which urges people to “do more” or “be better.” It also differs from the issue focus of most civic activism. Public work connects interests to citizenship and the public good by inviting people to “make work more public,” more interactive, collaborative, visible, and filled with public purposes.” (Boyte, 2014 p.2)

The ‘system world’ referred to by Boyte is implicitly set up against the ‘lifeworld’ (as distinguished in writings of Jürgen Habermas) of public work and civic agency. Whilst systems thinking often usefully incorporates ‘systems’ as ontological devices (real-world entities) as used by Habermas, the systems thinking in practice perspective adopted in this paper privileges ‘systems’ as epistemological devices (that is, rendered versions of real world entities for understanding and transforming realities). Systems are characterised here as being purposeful rather than purposive. Aside from the provisional character of ‘reference system’ in CSH where the underlying purpose (CSHq2) is continually subject to change, the idea of purposefulness in a CSH reference system is further captured with Fig.3 in relation to each source of influence associated with ‘developing’ rather than entrenching stakeholding positions (cf notions of developing values, developing power, developing know-how, and developing legitimacy). Reynolds (2014b) makes the point that evaluative work using CSH can be associated with the tradition of developmental evaluation (Patton, 2011).

A generic model of developmental evaluation based on systems thinking in practice (STiP) can help render the types of conversation required not only for (developmental) evaluation processes, but also for delineating the usefulness of evaluation as ‘public work’ in civic professionalism. Referring back to a CSH reference system (Fig.2), one of the key ingredients of any conversation on professionalising evaluation is a concern over competencies (CSHq8). Prevailing models of social trusteeship and technocratic professionalism have much to say on listing the various competencies associated with a professional evaluator. Surprisingly little is said regarding the competence of making judgements – presumably based on the premise that evaluators are meant to be dispassionate or objective and therefore ‘non-judgemental’: “In a study of evaluation practitioners
who rated the importance of 49 evaluator competencies, the competency of “making judgments” was rated the second lowest on average of all competencies, indicating that not all respondents believed that to evaluate is to judge” (Schwandt, 2016 p.9 citing King et al., 2001. my italics)

In contrast, this paper argues that making judgements is at the heart of evaluation practice and indeed should be the cornerstone of any civic professionalism. As noted above, the idea of contrasting factual judgements with value judgements through the medium of boundary judgements is implicit in the workings of CSH. Figure 4 depicts this process as part of a wider notion of praxis; using systems as integral to purposeful conversation (cf. Reynolds, 2014a; 2016).

The STiP framework comprises three entities - evaluands as ‘situations of interest’, evaluators as one amongst other stakeholders associated with civic agency, and evaluations as rendered ‘systems of interest’ - and three associated activities. These activities involve making evaluation judgements which can be associated with ideas of seeking worth, merit and significance (Scriven, 2011): that is, regarding (i) ‘worth’ as factual judgements associated with understanding inter-relationships; (ii) ‘merit’ as value judgements associated with engaging with perspectives; and (iii) ‘significance’ (boundary judgements associated with reflecting on boundaries. The praxis of STiP for use in evaluation requires keeping all three activities at play (cf. Reynolds, 2015a). Similarly, such praxis may be of use in furthering conversation on the use and usability (worth, merit and significance) of professionalism in evaluation.

5. Conclusion
At the 2017 ‘Sustainability in Turbulent Times’ conference referred to earlier, Professor Andy Stirling from the Science Policy Research Unit at University of Sussex, spoke of a distinction between realms of policy and politics in relation to the role of experts. The distinction aligns with CSH distinction
between those ‘involved’ with policy and those ‘affected’ by policy. Stirling argued that in the post-trust era of turbulence and uncertainty, experts (advisors, researchers, evaluators etc.) in order to (re)claim social legitimacy, need to grapple as much with the political realm as with the more conventional (and comfortable) policy realm.

This paper has sought to explore the usefulness of evaluation for supporting policy (or any other intervention including programmes and projects) and the politics of interventions, through the use and usability of (i) models of professionalism, and particularly the model of civic professionalism, and (ii) systems thinking in practice (STiP) and associated tools like critical systems heuristics (CSH) as a means of modelling (rendering) complex issues of professionalism.

Like evaluation practice itself, the conversation on professionalism requires a judgement call. From this paper’s perspective the jury remains out on whether the usefulness of evaluation might be enhanced through enacting professionalism. Whilst the jury maybe out, there is scope for enhancing the use and usability of evaluation as a profession when bounded by meaningful conversation on evaluation as ‘public work’ framed through civic professionalism. Using CSH models to render such ideas of civic professionalism in evaluation, as part of a wider modelling of STiP based on praxis (juxtaposing judgements of ‘fact’, with value judgements through boundary judgements), can in turn enhance the use and usability of evaluation as praxis based on an ethos of ‘public work’.

References


