



# Open Research Online

---

The Open University's repository of research publications and other research outputs

## Towards Systemic Evaluation in Turbulent Times – making a second-order practice shift

### Journal Item

How to cite:

Schmidt-Abbey, Barbara; Reynolds, Martin and Ison, Ray (2020). Towards Systemic Evaluation in Turbulent Times – making a second-order practice shift. *Evaluation*, 26(2) pp. 205–226.

For guidance on citations see [FAQs](#).

© [not recorded]

Version: Accepted Manuscript

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's [data policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](https://oro.open.ac.uk)

# *Evaluation*

## **Towards Systemic Evaluation in Turbulent Times – making a second-order practice shift**

Journal:	<i>Evaluation</i>
Manuscript ID	EVI-19-0046.R1
Manuscript Type:	Article
Topics of Interest:	Complexity and Systems, Evaluation Norms, Standards and Ethics, Evaluation Practice & Professionalisation , Values and valuing, Evaluation learning, influence and use
Approaches & Methods:	Formative, Participatory and Action Research approaches, Evaluation design and choosing methods, Theory building and Theory Use
Abstract:	<p>The need for, and possibilities of, a second-order shift in evaluation practice are explored. Second-order evaluation practice enables an evaluator to improve practice as a skilled practitioner, acknowledging her embeddedness within an evaluand. The paper explores evaluation practice as experienced by professional evaluators, using ideas from developmental evaluation coupled with systemic evaluation in the tradition of systems thinking in practice (STiP). Systemic evaluation aims to capture systemic sensibilities – the bigger picture – of complex turbulent situations of change underpinning evaluands. Attributes of second-order practice with systemic evaluation are understood as being aligned with both systemic and systematic modes of evaluation praxis.</p> <p>Personal experiences are provided where this juxtaposing praxis has been found wanting. By example, a STiP framework is explored as heuristic support for making a second-order practice shift. The paper concludes with a discussion of some implications for developments in professionalising evaluation practice and research.</p>

**SCHOLARONE™**  
Manuscripts

## Towards Systemic Evaluation in Turbulent Times – making a second-order practice shift

### 1. Introduction

Interest in systems thinking in evaluation has come and gone in cyclical waves over some decades. As Elliot Stern remarked in a 2011 editorial to this journal: “Systems thinking in the social sciences waxes and wanes and now appears to be once more in the ascendancy” (Stern, 2011: 324). Interest is sustained through conferences like the European Evaluation Society (EES), where topics devoted to systems approaches and complexity in evaluation have featured prominently over the last decade. The American sister organisation (American Evaluation Association or AEA) even has an active ‘Topical Interest Group’ on Systems in Evaluation.

In theory, if not always in practice, the evaluation community is aware of the many challenges encountered by evaluation practice within turbulent, complex situations of change and uncertainty. A call for more systems thinking has been a common response to these experiences along with an increasing sense of urgency to step-up the changes required in practices to more effectively address situational turbulence (Catwell and Sheikh, 2009; Piirainen et al., 2012; Reynolds et al., 2015, Caffrey and Munro, 2017). For example, the European Evaluation Society’s biennial conference 2020 (EES 2020) is devoted to ‘evaluation in an uncertain world’ and the importance of complexity, legitimacy and ethics. The previous conference (EES 2018) was dedicated to the search for evaluation for more resilient societies.

This paper is grounded in evaluator practitioner experiences where there is a gap between espoused systems thinking, or systems-thinking-in-theory, and systems-thinking-in-practice (STiT c.f STiP). The argument is made that a ‘second-order shift’ in practice by evaluation practitioners is needed to foster the emergence of *systemic evaluation*. In addition to practitioner experience, this proposition draws on ideas from developmental evaluation and the expanding tradition of Systems Thinking in Practice, or STiP (Ison 2017). We explore the central role the evaluation practitioner ought to play at the intersection of evaluation practice, the context and situation where the evaluation practice is enacted, and the choice and use of appropriate theory, methods and frameworks for systemic evaluation.

A systemic conception of practice and the practitioner is first introduced, followed by the role second-order practice can play in evaluation. We then introduce systemic evaluation realised through STiP, as a primary means of enacting second-order practice. How evaluators can act as reflexive, systems thinking practitioners is then explored through considering their own Being, Engaging, Contextualising and Managing. Following an exploration of the implications of our arguments for professional evaluation practice, we conclude by raising questions critical to the further development of second-order practice in evaluation. Key to such a shift, we

1  
2  
3 argue, is the use of a reframing within systemic evaluation practice leading to a focus on ‘small  
4 r’ research practice associated with the design and enactment of learning systems.  
5  
6  
7

## 8 **2. Second-order practice and practitioners**

9  
10  
11

12 Evaluation literature and practice, including evaluation professionalisation discourse, share a  
13 strong focus on theoretical frameworks, methods and tools. An element that is often  
14 neglected is the evaluation practitioner i.e., knowingly or not the practitioner is abstracted  
15 out of practice. Frequently, the evaluation practitioner is literally ‘not in the picture’ in  
16 methodological discussions about evaluation. Much of the attention in the literature and  
17 professional discourse is given to methods, which concern the ‘how’ of evaluation  
18 implementation. In these discussions, evaluation guidelines, standards and ‘best’ practices  
19 are typically geared towards conducting ‘robust’ evaluations.  
20  
21  
22  
23

24 For example, the Evidence-Based Policy-Making approach favours certain methods over  
25 others. Randomised Control Trials (RCTs) and other experimental approaches are often hailed  
26 as a ‘gold standard’. Whilst there undoubtedly are cases to be made where RCTs and  
27 experimentation can be used to good effect, the promotion of ‘gold standards’ and  
28 proliferation of ‘standards of evidence’ is the antithesis to complexity-sensitive or systemic  
29 evaluation. The evaluation and evidence-based policymaking communities have extensively  
30 debated these contentions in the recent decade: for example, Cairney (2016), Duffy (2017),  
31 French (2018) but the contestations are far from settled.<sup>1</sup>  
32  
33  
34  
35

36 Frequently, the impression is created that the desired robust evidence can be produced in an  
37 objectified manner, with an assumption of reproducibility, as if dealing with a clinical  
38 experiment. Rarely is the practitioner herself given much attention in these methodology-  
39 focused discussions.  
40  
41

42 A shift from first-order (exploring the world) to second-order (reflecting about the  
43 exploration) research centrally involves the role of the practitioner engaged in such  
44 explorations. In the dominant first-order research tradition, the researcher is an independent  
45 objective observer who is outside the situation of concern, which is treated as the object of  
46 research. The individual doing the observation is not of concern and assumed to be an  
47 objective and dispassionate (and replaceable) observer and investigator. In the evaluation  
48 field, this separation of the observer from the situation of interest is even further accentuated  
49 by the emphasis put on evaluator independence and impartiality, which implies being external  
50 to the evaluand.  
51  
52  
53  
54

---

55  
56  
57 <sup>1</sup> The readership of this journal is well informed about the content and contentions of this methodological  
58 discourse, which goes beyond the scope of this paper.  
59  
60

1  
2  
3 By contrast, the practitioner in a second-order research approach reflects upon their  
4 explorations as being integral to the situation of concern, of which he/she is part. Observer  
5 inclusion is a hallmark of second-order research approaches, such as second-order cybernetics  
6 in the traditions of Heinz von Förster (1984, 1992) and Humberto Maturana (Maturana and  
7 Poerkson, 2004).  
8  
9

10 Drawing on the traditions of systems thinking and second-order cybernetics, we propose that  
11 a second-order shift towards systemic evaluation requires more attention to be paid to what  
12 it is that an evaluation practitioner (evaluator or evaluation commissioner) actually does,  
13 when she *'does what she does'*? (Ison 2017: 5, paraphrasing Maturana).  
14  
15  
16

17 Praxeology is the study of human action, relating to engaging in purposeful and willed  
18 behaviour. Practice is often referred to in relation to a professional practice: it is what  
19 professionals 'do when they do what they do' – this defines a practitioner's practice (based  
20 on Ison 2017: 14). There is also praxis, which concerns a theory or philosophy becoming a  
21 practical social action.  
22  
23  
24

25 What do evaluation practitioners actually do as their practice? The default assumption of what  
26 may be meant by an evaluation practitioner is being an evaluator, conducting evaluations. But  
27 it can involve any other practice role in relation to evaluation, such as: evaluation  
28 commissioner, evaluation user, or any other stakeholder involved and affected by an  
29 evaluation, or evaluation researcher. As Wadsworth (1997) noted we all do evaluation every  
30 day.  
31  
32  
33

34 The evaluation tradition of Developmental Evaluation (DE) invokes the idea of the evaluator  
35 as being embedded in the evaluand (Patton, 2011, 2018). Our paper shares this view, with  
36 the difference that (i) we explicitly call this second-order practice, and (ii) make it relevant to  
37 all evaluation practices in all evaluands – not contingent on specific circumstances of the  
38 evaluand or situation.  
39  
40  
41

42 One of the hallmarks of becoming a skilled practitioner (in any field) is being a reflective  
43 practitioner. The key concepts introduced by Donald Schön in the 1980s (Schön, 1984, 1987)  
44 are reflection-on-action and reflection-in-action. The former is ex-post whilst the latter is  
45 enacted and thus embodied in the unfolding doing of practice. The reflective practitioner  
46 concepts are integral to second-order practice; they imply that the practitioner reflects-on  
47 and –in practice about their practice, including their language, assumptions, values,  
48 repertoires, theories and emotions. Ison (2017) explains reflexivity as a reflection-on-  
49 reflection, a second-order practice that encompasses both of Schön's distinctions.  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## 2.1. What is meant by a 'second-order shift'?

Grounded in empirical work with pastoralists in semi-arid Australia, Ison (2017: 278-282) makes a distinction between 'first-order' and 'second-order' research traditions (see Ison and Russell 2000). The 'first-order research tradition' continues to dominate how science and research – and evaluation – are practiced. A first-order research tradition – applied to the field of evaluation - is characterised by the dominance of well-established social science research approaches based on notions of linear causation, a systematic linear chain of causal factors. The purpose of an evaluation in first-order understanding is to systematically observe and record such chains with appropriate levels of dispassionate objectivity. Ideas of linear causation are typically expressed for instance in traditional log frame models, by targets, goals, and the understanding of objects of study or evaluation as fixed entities that can be studied and measured objectively. An expression of first-order evaluation practice might be associated with the founding traditions of evaluation as a discipline through the works of Scriven (1991, 1996, 2001, 2003), features of which have been described by Patton as 'external accountability' (Patton, 1994: 318). A shift from a first- to second-order tradition involves a shift in focus from researcher/evaluator objectivity to researcher/evaluator responsibility.

Methodologically, it is regarded as good practice within the evaluation field to draw on the methods from a repertoire of well-established social science research, with a preference for multiple methods, and methodological plurality, as defined in evaluation guidelines of major evaluation professional organisations at national and international level (e.g., the AEA and the European Evaluation Society or EES). Arguably most guidelines support evaluation practice in terms of first-order, external accountability.

A second-order research tradition on the other hand is characterised by experiential and relational understanding of practitioners engaging with situations they themselves are part of rather than being distanced observers. In evaluation practice, this would for example involve critical reflections and reflexivity about a reality that includes the evaluator. The reality is brought forth with participation and inclusion of the evaluator and other practitioners involved in the situation.

Second-order reflexivity has found its way into some forms of evaluation practice already. For example, through the incorporation of elements of action research, as practicing evaluation as an ordinary everyday part of what we do. Wadsworth (2016) introduces doing 'evaluation on the run' to non-specialist evaluators, to develop a culture of ongoing evaluation as part of their normal business. This bears some common features with internal evaluation (Love, 1991), and enacts continual action-learning cycles of observation, reflection, dialogue and implementation to be applied to all our actions (a form of small r research) as cycles of continuous monitoring and evaluation (Wadsworth (2016: 94).

1  
2  
3 Similarities might also be found through some aspects of DE (Patton, 1998, 2011). Patton  
4 comes close to describing second-order practice in relation to DE, whilst not using this term  
5 or overtly drawing upon its associated intellectual and methodological lineages (Patton, 2018).  
6  
7

8 Some common features within the STiP tradition suggest that DE and Blue Marble Evaluation  
9 might provide a promising avenue towards systemic / second-order research and evaluation.  
10 However, DE to date does not make explicit claims towards advocating second-order practice.  
11 Furthermore, DE is founded on a contingency viewpoint; suggesting that DE is only  
12 appropriate for specific evaluands, such as niche interventions associated with social  
13 innovation.  
14  
15

16  
17 DE also considers the evaluator as an embedded and embodied constituent stakeholder along  
18 with others in the evaluand. In situations or evaluands deemed by DE practitioners as being  
19 complex, the role of the evaluator requires a sense of ethical internal responsibility where the  
20 evaluator is part of, rather than external to, the evaluand. But what is complex is not an  
21 ontological choice but an epistemic choice based on the capabilities enacted through practice;  
22 as Cook and Wagenaar (2012:9) claim in their epistemology of practice “as an inquiry into the  
23 possibilities and constraints of being engaged, embodied, contextualized agents.”  
24  
25  
26

27 Our claim in this paper, in contrast, is that second-order evaluation based on a tradition of  
28 STiP, has value in all evaluands, amongst all evaluators, and for all evaluations (see Reynolds  
29 2015).  
30  
31

32 Features of a critically reflexive systemic evaluation that would constitute a second-order  
33 research tradition would include, for example:  
34

- 35 • Appreciating that the evaluand’s context is constantly changing.
- 36 • Questioning the terms of reference for an evaluation (the ascribed purposes/standards  
37 used).
- 38 • Iterating on measures of success and other criteria used for evaluation;
- 39 • Adapting tools at hand rather than seeking a reified ‘best-practice’ or ‘best-fit’ tool(s).
- 40 • Regarding the evaluator as part of the evaluand rather than separate from it.

41  
42 In terms of desired outcomes and benefits, it is appropriate to recall three principles for more  
43 systemic evaluation, initially proposed by Reynolds et al. (2015). We argue that through the  
44 enactment of such a second-order shift, evaluations, evaluators and the evaluands can  
45 display:-  
46  
47  
48  
49

- 50 (i) more systemic, reflexive and humble boundary conversations between values  
51 (evaluations) and unbounded reality (evaluand);
- 52  
53 (ii) more empathic, ethical and response-able engagement with evaluand  
54 stakeholders based on reflexivity; and  
55  
56  
57  
58  
59  
60



(iii) a more adaptive use of 'tools' and methods as part of evaluation praxis, whilst recognising the limitations and ultimate fallibility through increased epistemic awareness.

Other examples of evaluation initiatives implicitly aligned with second-order practice include, as mentioned, Wadsworth's (2016) 'evaluation on the run' to develop a culture of ongoing evaluation as part of normal business, values-based evaluation (Hall et al., 2012), and Schwandt (2017) calling for more democratic professionalism in evaluation. A specific interest for systems approaches in evaluation lies in the area surrounding values and ethics, and the use of boundary critiques, c.f. works by Schwandt (2015, 2017, 2018), and Schwandt and Gates (2016). Recently, Schwandt (2019) has signalled the emergence of 'post-normal evaluation': mirroring the now established discourse of 'post-normal science' (Funtowicz and Ravetz, 1993).

The second-order systems approach advocated here situates the practitioner as central to their own practice and thus moves the debate from what is the best method or approach to what might be the best enactment, or performance (in the sense of a choreographer or dramaturgist), of contextualised systemic evaluation. The following vignette exemplifies what is at issue.

## 2.2. Vignette 1: Experience of a 'jobbing evaluator' working in first-order practice

The lead author works as an internal evaluation practitioner in an organisation. This role can be described as a 'jobbing evaluator', whose main professional responsibility is to engage professionally with evaluations.

Such a professional evaluation role is distinct from what might be described as a crafting / bricoleur evaluator: a practitioner who embeds evaluation into other professional practices through creative application of evaluative thinking and acting.

The day-to-day experience of working as a 'jobbing evaluator' involves first-order evaluation practices primarily aiming at external accountability. In the regulatory framework and context of the organisation, evaluation of activities and programmes is a regulatory requirement. Therefore, evaluations 'must be done', and are subject to reporting and auditing. Great emphasis is put on evaluator independence, and summative and formative evaluations, although evaluation is also valued as a source of organisation learning.

Standard professional tasks of this 'jobbing' evaluator include developing an evaluation policy and programme, developing terms of reference, commissioning evaluations to be conducted by external evaluation contractors, contract management, liaising with evaluation stakeholders, and ensuring evaluation findings and recommendations are useable, used and acted upon, and feed into organisational learning and development, as well as conducting, where appropriate, some evaluations directly as an internal evaluator. These are typical tasks shared by many monitoring, evaluation and learning (MEL) practitioner roles in many organisations, involving systematic application of professional evaluation good-practices. An outcome of the systematic approach is the production of standardized guidelines, sometimes blueprints and protocols rather than the creation and recreation of a contextualised systemic evaluation performance.

Over time I developed a more critical appreciation of the nature of the information and knowledge provision for policy and decision-making purposes in complex situations using this standard approach. I became increasingly uncomfortable with the available methodologies and the evidential claims derived from them.



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Through the study and application of STiP, I learnt that there are different ways to frame, study and evaluate interventions and policies. I came to understand that providing knowledge and evidence to policymakers is steeped in positivist assumptions and favours linear knowledge transfer models which constrains opportunities to introduce a more systemic approach to evaluation in practice.

As the experience in vignette 1 shows, a jobbing evaluator's practice can be firmly – or even exclusively - rooted in first-order practice, often to the perhaps unconscious exclusion of second-order practice. A concern we have is how STiP can provide the means for creative expansion of first-order evaluation practice, towards second-order evaluation practice.

### 3. Systemic evaluation informed by Systems Thinking in Practice (STiP)

Systemic evaluation (SE) can be considered as part of a wider tradition of systems-based evaluation. SE might be a means to correct over-systematic features of conventional first-order practice. To explore the value of systems thinking further, it is necessary to first appreciate usage of the systems idea derived revealed from within the STiP tradition. We offer distinctions between conventional systems-based evaluation (based on STiT), and systemic evaluations (based on STiP).

a. Systems-based evaluation is largely systematic-oriented evaluation in a first-order tradition, which is looking at the 'system' understood as real, ontological devices, following a positivist worldview. This worldview may be held with awareness or unknowingly. In this understanding, a system is considered to exist as a real entity, to be systematically studied through scientific methods including modelling and characterisation. Systematic systems-based evaluations are afforded by use of the log frame; evidence-based evaluations and use of experiments and methods such as Randomised Control Trials.

b. In contrast to this ontological understanding of systems-based evaluation is systemic evaluation in the tradition of STiP, rooted in a constructivist understanding of systems. In this understanding, systems are brought forth, or distinguished by, practitioners interested to engage with a situation and understand or change it systemically. The role of a system is to be used as an epistemological rather than ontological device i.e., as a way of knowing about a situation of concern, including an evaluand.

c. With systemic sensibility, systems literacy and STiP capability (Ison and Straw 2020) the systematic and the systemic can be regarded as a duality, as combining to constitute a holistic response to a situation of concern. Reframing as a duality enables breaking away from the domination of systematic thinking and practice without

1  
2  
3 abandoning it, as happens when unhelpful, self-negating pairs are constituted as  
4 dualisms (e.g., mind – body). Awareness and internalisation of these distinctions  
5 enables a productive dynamic between systemic and systematic practice – frequently  
6 switching between them as required in a dynamic, interdependent, contextual and  
7 emergent relationship.  
8  
9

10  
11 The distinction between an ontological and epistemological understanding of systems is  
12 important here. An inherent risk of naively mapping systems in comprehensive systems maps  
13 – as ontological devices - is the temptation to confuse the ‘map’ with the ‘territory’ (Korzybski  
14 1933: 58). A systems map developed without STiP capability can convey the misleading illusion  
15 that ‘this is the system, and it shows it the way it is’. But it is important to consider that any  
16 systems map is only a – still partial, and biased – representation of what is perceived to be the  
17 system by stakeholder-practitioners, including modellers. A systems map is how it appears to  
18 the modellers from their respective lenses, which is influenced by many factors and subject to  
19 biases, partial perceptions and omissions. "All models are wrong" (Box, 1976: 792) also applies  
20 to systems models - keeping in mind that they may still be useful, with the necessary caution  
21 and epistemic awareness of their limitations.  
22  
23  
24  
25

26 As we will discuss, the bringing forth of systems as epistemological devices raises ethical and  
27 design concerns and possibilities. For a system to be a system involves someone making a  
28 boundary choice – a distinction between what is in and what is out of a system-of-interest.  
29 Applied to systemic evaluation and to overcome this difficulty, we therefore offer the  
30 distinction that evaluands are *situations of interest* to be explored by the evaluation, and not  
31 ‘systems’ as such.  
32  
33

34 Over the last decade and longer, authors have advocated a need for the evaluation profession  
35 to engage with systems thinking and complexity science (STCS), for example, Patton and  
36 colleagues (2007, 2011), Williams and Imam (eds) (2007), Midgley (2007). Practitioners’  
37 toolkits were provided by Williams and Hummelbrunner (2009), and Reynolds and Holwell  
38 (eds) (2020) offer a guide to systems approaches to change from a STiP tradition.  
39  
40  
41

42 Of concern to us is the extent to which these claims are becoming institutionalised within the  
43 evaluation community and whether they are being institutionalised as first, or second-order  
44 STiP praxis, or both. It is apparent that investment in greater STiP capability combined with  
45 conducive institutional innovation that allows STiP to flourish are warranted.  
46  
47  
48  
49

### 50 51 3.1. Three elements for systemic evaluation: interrelationships, multiple perspectives 52 and boundaries – applied to evaluation practice 53 54

55  
56 From epistemic awareness in STiP it follows that three elements of how systems can be viewed  
57 and approached analytically are particularly key for evaluation (c.f. Reynolds, Holwell, [2020],  
58 Williams, 2013):  
59  
60

- 1
- 2
- 3 • *Interrelationships* between systems components;
- 4 • *Multiple perspectives* through which a system can be viewed, by different
- 5 stakeholders and from different worldviews;
- 6
- 7 • *Boundaries*: the way judgements are made about what is 'in' the system-of-
- 8 interest, and what is 'out', and a critical engagement with these boundary
- 9 judgements.
- 10
- 11

12 These three core concepts find resonance in methods and approaches developed in the STiP  
13 tradition which enable evaluations in practice i.e., to building systemic evaluation capability.  
14 Exploring situations regarded as evaluands through the lens of these three concepts opens up  
15 innovation-through-design possibilities as well as enabling ethically defensible praxis.  
16

17  
18 Fortunately, showing the *interrelationships, interdependencies and causal links* between  
19 systems components has gained interest in the evaluation field in recent times. There is  
20 growing interest to 'show the system to itself' through the use of visual representations, or  
21 systems maps. Undertaken with epistemological awareness STiP practitioners realise how  
22 understanding (learning) can be enhanced by choosing to map elements in a situation-of-  
23 interest *as if* they were a system. Mapping systems visually has the intention of seeking to  
24 understand and representing system components, causal links and behaviours as  
25 comprehensively and succinctly as possible. Systems maps can consist of diagrams created  
26 through a variety of systems mapping methods (Blackmore et al 2017). Epistemologically  
27 aware practitioners always carry at the forefront of their practice the questions: Whose  
28 system? Whose boundary judgments? Their practice can also reveal whether what might be  
29 perceived as a system actually functions as a system and what its purpose may be from the  
30 perspective of different stakeholders.  
31

32  
33 A well-known example is the obesity systems map (UK Government Office for Science, 2007).  
34 This diagram captures a multitude of factors and interrelationships in one picture. It illustrates  
35 the simultaneous strength and weakness of such diagrams: there can be a temptation to pack  
36 'everything in' to be comprehensive and in trying to capture 'the whole system'. This ambition  
37 comes at the cost of understandability and accessibility for the reader. At first sight, such a  
38 diagram can be casually described as a 'spaghetti diagram' and is very difficult to read. It may  
39 have the opposite effect than what is intended, as it can put a reader (for example,  
40 policymaker) in a position of feeling overwhelmed by the complexity and lead them to  
41 disengage, rather than feeling empowered to see intervention possibilities.  
42

43  
44 In the hands of an experienced STiP practitioner the different choreographic possibilities of  
45 their practice with other stakeholders are appreciated. A systems map done alone is not the  
46 same as one done with others (raising the question of which others). The act of mapping as  
47 an emerging creation, following multiple iterations with others, is in itself a mini-learning  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 system<sup>2</sup>. Primary insights arise in the process itself among those participating. A final map,  
4 used for presentational purposes, creates far more limitations in comparison to an enacted  
5 learning system. What is systemic often becomes systematic, following a linear mode of  
6 communication-as-delivery.  
7  
8

9 *Multiple perspectives* of a system-of-interest by different stakeholders are traditionally in  
10 focus in evaluations. Many evaluation approaches have been developed to address and  
11 capture different stakeholder perspectives (for example, participatory evaluation,  
12 empowerment evaluation and other approaches). Systems approaches can be combined and  
13 supplement these well-established evaluation approaches. For example, Soft Systems  
14 Methodology (SSM) is well suited to dig deeper into the different perspectives – not only  
15 descriptively – what different stakeholders think – but also what the underlying philosophical  
16 foundations are that influence these differences in perspectives: the underlying ‘worldviews’  
17 (Weltanschauungen) which inform the different perspectives (Checkland and Scholes 1990).  
18  
19

20 *Making boundary choices and judgements*, and critical reflections about these choices is  
21 central to evaluation design and implementation. For example, evaluation commissioners  
22 frequently pre-determine the boundaries of what is in scope for an evaluation in the  
23 evaluation terms of reference (ToR). With a view to the stakeholders that are affected but not  
24 necessarily involved in an evaluation, it is a question of evaluator responsibility (ethicality) to  
25 be critically aware of the boundary choice that is implied in the scope of a ToR. It is therefore  
26 important for evaluators to have ways to ethically address these choices and their  
27 implications. As an example, Critical Systems Heuristics (CSH) (Ulrich 2010) can be mentioned  
28 as one of the more popular systems approaches with application in evaluation practice,  
29 possibly for this reason. CSH has been found useful and has been used by evaluators and quite  
30 extensively written about. Examples can be found in works by Ulrich and Reynolds (2020),  
31 Gates (2017), and Stephens et.al (2018).  
32  
33

34 CSH is well suited to conduct boundary explorations and critiques and can also be helpful in  
35 revealing inter-relationships and multiple perspectives. Through the format and underlying  
36 concepts, CSH can offer a ‘framework for understanding’ a situation being evaluated. CSH  
37 contains twelve critical questions that can be asked about a situation, in two different modes:  
38 how it ‘ought to be’ (in an ideal situation), and how it ‘is’ (in reality). This contrast between  
39 normative and actual mode of application of these questions can be used dialectically to  
40 critically explore and expose the boundary choices that have been made and highlight the  
41 consequences of these boundary decisions across different dimensions.  
42  
43

44 From a STiP perspective the task of exploring an evaluand is always to start systemically (as  
45 with Reynold’s model of an evaluation-adaptive complex (Reynolds, 2015). This involves not  
46  
47  
48  
49  
50  
51

---

52  
53  
54  
55  
56  
57 <sup>2</sup> A ‘learning system’ has “built-in capacity for reflection on experience and to recognize change or learning  
58 when it occurs” (Ison and Russell, 2000:208)  
59  
60

1  
2  
3 (systematically) pre-judging a situation as being either simple, complicated, complex, or  
4 wicked, but rather to assume that the situation will in all likelihood have a mixture of features  
5 which vary with perspective i.e., the situation of concern will be open to the framing choices  
6 of the practitioner(s). STiP is then about making the situation more amenable to purposeful  
7 action by bounding the evaluand within a situation of concern, but always with some element  
8 of systemic awareness.  
9  
10  
11  
12  
13

### 14 3.2. Complexity thinking in practice 15 16 17

18 Complexity thinking – which is (arguably) a sub-set of systems sciences - has come into good  
19 currency in the evaluation profession over the last decade, as can be seen by numerous  
20 publications and discursive discussions at conferences. Forss et al.'s (2012) compilation of  
21 'evaluating the complex' was one attempt to bring complexity science concepts into view in  
22 evaluation theory and practice. Other authors like Ramalingam (2013) and Bamberger et al.  
23 (2016) have explored how dealing with complexity can be applied to the domain of  
24 development evaluation to become more complexity responsive.  
25  
26  
27

28 Tensions between systems thinking and complexity science advocates in the evaluation field  
29 may appear to have been resolved by detecting an emerging relationship between complexity,  
30 systems thinking and evaluation (c.f. Reynolds, Forss et al., 2012). The debates continue  
31 though as documented in special journal editions e.g., special editions of the Bulletin of the  
32 Institute of Development Studies (IDS, 2014) dedicated to 'exploring the potential of systems  
33 ideas and complexity concepts to meet the increasingly complex challenges of an increasingly  
34 ambitious development agenda' (p. 1). The tensions exposed and debated between systems  
35 thinking on the one hand and complexity science on the other include contrasting uses of the  
36 systems idea as either an ontological device for understanding complex systems i.e., modelled  
37 as real observable entities, or alternatively as epistemological devices for understanding and  
38 engaging with situations i.e., as constructions for social learning and understanding. Within  
39 the complexity discourse 'complex adaptive systems' (CAS) is frequently used unreflexively as  
40 an ontological device rather than as a conceptual framing which can be chosen; i.e., what  
41 could be gained by considering this situation *as if* it were a CAS?  
42  
43  
44  
45  
46  
47

48 Despite recent investments and some promising examples there continue to be relatively few  
49 cases of genuinely systemic evaluations in practice, although this may be changing. Kusters et  
50 al. (2019, p. 34) list several promising examples. With reference to monitoring and evaluating  
51 the SDGs, IIED (2019) has called for 'complex systems thinking and using systemic approaches  
52 to evaluations that connect'. The latter is also reported by Ofir et al. (2019). There are some  
53 examples in the areas of eco-systems (Mueller and Sukhdev, 2018), and examples in the works  
54 of Stephens et al. 2018.  
55  
56  
57  
58  
59  
60

1  
2  
3 Despite growing calls for systemic evaluation, there continues to be a gap between what is  
4 claimed and promoted (namely, to promote whole-systems evaluations or complexity-  
5 responsive evaluations) on the one hand, and the actual implementation and practical use of  
6 systems and complexity approaches and methods in day-to-day evaluation practices, outside  
7 of the show-cased examples in the literature.  
8  
9

### 10 11 12 13 3.3. The research-for-practice-reform gap 14 15

16  
17 Existence of a practice gap resonates with the lead author's own research and professional  
18 experience based on attempts to introduce systems approaches in evaluation practice 'by  
19 stealth'. This experience motivates a programme of 'small r' and 'big R' research to address  
20 this gap based on first-person inquiry and case study research respectively. 'Stealth' attempts  
21 expose practical difficulties and tensions. Documented examples of these difficulties were  
22 reported in a 2019 workshop report by the 'Scaling Solutions toward Shifting Systems  
23 initiative' of the Rockefeller Philanthropy Advisors, devoted to assessing systems change, and  
24 building philanthropic funding organisations' capacity. Workshop participants were asked to  
25 identify what they believed were the barriers to adopting systems evaluation approaches in  
26 their organisations and the [development] sector as a whole. A diverse set of observed  
27 barriers was elicited (Rockefeller Philanthropy Advisors, 2019: 2). This list ranges from lack of  
28 knowledge and appreciation, unclear definitions and concepts to lack of resources and  
29 organisational capacities, and the dominant use of the logic-model paradigm for evaluations  
30 which is not suited to systems-wide change.  
31  
32  
33  
34  
35  
36

37 In addition to this list, from the authors' own experiences in the field, the scarcity of practical  
38 examples of systemic evaluations and reported difficulties point to a gap between an  
39 espoused theory of what is claimed needs to be done (for example, to conduct an evaluation  
40 that does justice to the complexities encountered in the evaluand) and what happens in  
41 practice (theory-in-use) (Argyris and Schön, 1974).  
42  
43  
44

45 This poses the question: how can such a gap be bridged? The primary argument arising from  
46 our work is that taking a second-order approach to evaluation practice can help to overcome  
47 existing dichotomies and dualisms, to bridge the practice gap. A shift needs to happen, and  
48 yet there are many barriers along the way.  
49  
50

51 One such barrier is the epistemic contrast of applying systems thinking as a first-order  
52 (ontologically fixed) application compared to the second-order understanding of STIP as an  
53 epistemic approach for learning and exploration of a situation-of-interest. Vignette 2  
54 illustrates an experience by the lead author where an epistemic clash arose during the attempt  
55 to design a systemic evaluation. In this example, this contrast manifested itself as a dualism –  
56  
57  
58  
59  
60



rather than a duality – with the result that the ambition for a systemic evaluation was abandoned as the differences could not be reconciled.

### 3.4. Vignette 2 – Example of an experience of the gap between first-order application of systems thinking with second-order understanding of systems practice

The lead author was one of two OU STiP practitioners who were approached to get involved in an evaluation project which had an explicit ambition to incorporate systems thinking approaches.

The client had previous exposure to System Dynamics (SD), one of many theoretical and practical lineages within the systems field, and had a very positive experience of her applications of SD. From the perspective of this practitioner explicit inclusion of SD into the project design was needed in order to make the evaluation more systemic. The client's expectation was that the use of the SD technique of Causal Loop Diagrams (CLDs) in particular would enhance the understanding of the complexity of the situation to be evaluated.

Due to this explicit request for SD, the STiP practitioners felt they needed to secure some additional SD modelling expertise beyond their own for this project. They contacted an SD consultant with view to exploring the feasibility to form a joint project team. The members of this potential project team attempted to co-design a customised methodology for this project involving evaluation concepts such as theory-of-change, combined with elements of systems approaches consisting of Soft Systems Methodology (SSM) and Critical Systems Heuristics (CSH), and SD.

The experience of co-designing and negotiating this methodology turned out to be very difficult due to very different understandings of the systems traditions which were exposed during the exploration phase.

Over the course of the exploration, it emerged that the assumption of the SD consultant was that the situation to be modelled is a fixed system, which needed to be modelled as a whole using a range of interconnected CLDs. This understanding can be described as 'this is the system which I need to understand in order to engineer and model it'. The SD consultant understood their role as an external objective observer and expert modeller, who needed to objectively model 'the' system, in order to mirror it back to the stakeholders in the situation. [This understanding is similar to that of mainstream consultants, business analysts as well as some evaluator practitioners].

The STiP practitioners on the other hand had a very different understanding of the situation arising from their exposure to a different set of systems traditions: the 'system' concerned in the situation to be evaluated does not exist as such in a pre-conceived form. It is not a system 'out there'. It can depend on the different stakeholders involved who have very different experiences of the situation being evaluated (multiple perspectives), and for whom the evaluand may have very different purposes. Therefore there can be confusion, contestation and complexity. The boundaries of the evaluand represented as a system (what is considered to be part of the system or outside of it) also differ and can change – different boundaries can be drawn for different purposes and by different actors.

Rather than modelling some pre-conceived system, understanding the evaluand means to organise an exploration (or evaluation) of it as a learning system. The role of the STiP practitioner is then to organise a systemic inquiry for an exploration of the situation for the purpose of learning and transforming the situation of concern.

How did the experience end?

In this case example, the very different assumptions and understandings of systems that were exposed through this feasibility exploration were experienced as profound, with the result that the project was not feasible to be implemented in this combination, and leading to abandonment.

#### **Distinctions between first-and second-order traditions**



1  
2  
3 The example in vignette 1 shows how different practitioners grounded in different traditions  
4 of systems practice approach their practice in very different ways. It is well known that  
5 paradigmatic and epistemological commitments differ within disciplinary fields; the challenge  
6 is to bring them into awareness and conversation.  
7  
8

9  
10 In vignette 2 reflections about this practice experience revealed that the SD practitioner had  
11 approached the project from the understanding of a first-order systems tradition, whereas  
12 the STiP practitioners came to it from a second-order tradition. These paradigmatic  
13 differences in understanding needed mutual recognition and a shared language and  
14 repertoire and were not recognised from the outset. Initially, all members of the team had  
15 assumed that they were all systems practitioners – albeit from different schools - and it would  
16 therefore be possible to come to a shared understanding. The differences between first- and  
17 second-order systems traditions had been unexpressed and thus underestimated by both  
18 parties. Only through the failure of this project was the depth of the gap between the two  
19 understandings and epistemological stances revealed.  
20  
21  
22

23 The experience presented in vignette 2 reinforced the question of how a shift between first-  
24 and second-order systems traditions can be enacted between practitioners in given situations.  
25 The role of the practitioner him/herself within the dynamic between the practitioner(s), the  
26 situation, and the methods and frameworks used seems to be crucial, as well as the  
27 practitioners' relationships between themselves. Capacity of team members to have a  
28 generative conversation and joint reflection from which something new can arise from an  
29 experience seems essential i.e., reflexivity.  
30  
31  
32

33 In the following, we propose a way to create opportunities to bridge this existing gap, towards  
34 more systemic evaluation practices by putting evaluation systems practice and the  
35 practitioner at the centre (enacting small r research/learning) and proposing a research  
36 agenda (which can be understood more as big R research).  
37  
38  
39  
40

#### 41 **4. Evaluators as systems thinking practitioners?** 42 43

44 To make the desired shift towards second-order evaluation practice, evaluation practitioners  
45 can benefit from approaching their practice by drawing on heuristics that reveal key  
46 choreographic, or performative, relational dynamics. Drawing on the repertoire from the STiP  
47 tradition, a possible vehicle for enacting second-order practice that has been found useful  
48 within Open University (UK) STiP education is that of systems practice as comparable to a  
49 juggling act (Ison and Blackmore 2014).  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

#### 4.1. Systems practice as a juggling act for evaluation practitioners – a social dynamic

Ison (2017) conceives systems thinking and practice as an active social dynamic. To bring this to life, he introduces the isophor<sup>3</sup> of a systems practitioner as a juggler, keeping several balls in the air as part of the juggling act (p. 60). We here briefly apply the image (isophor) of this juggling act to the practice of enacting an evaluation.

The evaluation practitioner is the key player in the 'performance' of an evaluation. Similar to the juggler in a juggling performance, the practitioner is invited to 'juggle' four different 'balls' that need to be played and kept in the air during an evaluation (performance): these 'balls' are: Being, Engaging, Contextualising and Managing (BECM).

- *The 'B' ball*: this concerns the 'Being' of an evaluation practitioner, with awareness of his/her tradition of understanding that informs his/her evaluation practice, and the ethical responsibility he/she needs to take in enacting the evaluation. Being a reflective/reflexive practitioner is a key capability to nurture this capacity, to take ethical responsibility for our actions as evaluators. This concern extends to those that are potentially affected by the consequences of the evaluation, or the situation being evaluated. As evaluation practitioners, we need to be constantly aware of such consequences which instils our ethical responsibility.
- *The 'E' ball* is for 'Engaging' with the situation the evaluation is concerned about and engaging with the evaluand. Juggling the E-ball requires awareness and agency to make the choices available to the evaluation practitioner of how to engage with the real-life situation of an evaluation. How situations are perceived has important implications for the frames chosen for it and choices made for their evaluation. For example, a situation may be perceived as well-defined (normal), or uncertain and complex. Critically exploring and reframing our perceptions of the situations we engage with in evaluation and developing our appreciative settings (Vickers, 1970) is then key to hone the capacity to engage reflexively in situations we are evaluating (e.g., Ison 2018).
- *The 'C' ball* is about 'Contextualising'. Applied to evaluation, this concerns the choices of methods, techniques and tools to be used for an evaluation. When evaluation practitioners decide on which methods or tools they will use for an evaluation, they contextualise their evaluation practice to a specific evaluation situation, for example,

---

<sup>3</sup> Juggling in this case is not a metaphor (which takes one concept into a new domain) but an isophor in which the same dynamics are apparent in different domains. This follows the understanding of Maturana, where juggling is an isophor of the vision one wants to have to claim that one understands, for example, a biological or a cultural happening, such as effective systems practice' (Ison, 2017: 61).

1  
2  
3 for a specific evaluation assignment. In the STiP tradition of understanding, the  
4 distinction is made between tools, techniques, methods and methodologies. Ison  
5 (2017:167) describes methodology as the 'conscious braiding together of theory and  
6 practice in a given situation, as a context specific enactment'. It requires a broad  
7 awareness of concepts, knowledge of techniques and tools, and methods. The  
8 methodology involves the design of a combination ('bricolage') of adapted methods  
9 customised to fit for the specific situation, in a way that feels systemically 'right'.  
10 Contextualising may also involve the exploration of the evaluation purpose from the  
11 perspective of different stakeholders and stakeholdings which may differ from the  
12 overtly stated purpose in the ToR provided by the evaluation commissioners. A  
13 boundary critique could explore the contestations of the purpose of the system-of -  
14 interest, and the interests involved. An adaptation of the CSH twelve question  
15 framework (Ulrich, 1983, 1996) for example can lend themselves as bridging practice  
16 (between Contextualising and Engaging) to explore who is/ought to be the systems'  
17 client, or whose interests are / ought to be served by it (Ison 2017: 166).  
18  
19  
20  
21  
22  
23  
24  
25

- 26 • Finally, 'Managing' involves simultaneous looking both outside to interaction with  
27 others, as well as looking inwards. In evaluation practice, this concerns how the  
28 evaluation practitioner manages his/her involvement in the evaluation: the  
29 performance and the relationship with the evaluation commissioners and other  
30 stakeholders, and indeed the overall context. Effective management of an evaluation  
31 (juggling) performance crucially involves nurturing and maintaining of meaningful and  
32 engaging relationships through the flux of time (Vickers, 1978: 71-72).  
33  
34  
35

36 Centrally, the juggler herself is the person who keeps all these ball in the air, through their  
37 practice. Juggling involves the whole body and mind. The juggler uses her own body  
38 throughout the performance and thus brings it into being: by throwing the balls into the air  
39 and catching them again, balancing her body through contact with the floor in response to the  
40 motion of the different balls. It is a co-evolving and adaptive practice, an interactive social  
41 dynamic, in which the juggler is coupled in relation to the four balls.  
42  
43  
44

45 These BECM dimensions can unfold their full power and relevance by relating them to the  
46 core theme of STiP of juxtaposing systematic (or first order) and systemic (or second-order)  
47 practice, applied to evaluation practice.  
48  
49

- 50 • **'Being':**
  - 51 ○ expressed systematically can mean to apply the evaluation 'tools of the trade'
  - 52 instrumentally (to get the job done). The evaluator acts like a tradesperson.
  - 53 ○ In systemic mode, 'Being' involves ethics, by acknowledging wider consequences of the
  - 54 work done (evaluator as craft artisan/bricoleur).  
55
- 56 • **'Engaging':**  
57  
58  
59  
60

- 1  
2  
3  
4  
5  
6  
7  
8  
9
- in systematic mode, the evaluand is often framed as a tame (solvable) difficulty or problem. The evaluator acts with confidence amongst stakeholders, including commissioners and intended beneficiaries and co-evaluators.
  - In systemic mode, the evaluand is better framed as a (potentially) wicked problem situation.
- **'Contextualising':**
    - In a systematic approach, this means fulfilling contractual obligations, and keeping commissioners assured.
    - Understanding contextualising in systemic mode, it can mean to (courageously) question the terms of reference from the commissioner.
  - **'Managing':**
    - when viewed from a systematic framing, managing an evaluation focuses on maintaining immediate task-oriented relationships with stakeholders.
    - Approached from a systemic framing, managing an evaluation involves a concern for forging longer-term and wider relations, beyond the immediate task and concern.
    - When enacted as duality (i.e., both systematically and systemically) an evaluation can function effectively (i.e., be managed) by generating learning through feedback, and thus adaptation and change.
- 10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33

#### 4.2. Adding capacity building for second-order evaluation practice to the evaluation professionalisation agenda

Enacting the juggling of these four balls can deliver one or both of systematic effectiveness or systemic effectiveness. Its use shows how it is possible to re-frame the engagement and for doing systemic evaluation by switching between the two modes, as appropriate to the situation. Von Förster (1992) understood this as the essence of an ethical performance because more choices are offered to those within the situation.

The STiP focus on the evaluator as practitioner and their central role in enacting responsible and systemically desirable evaluations could be seen as a response to the call for an enhanced evaluation ethos within evaluation professionalisation discourse. Schwandt (2015, 2017) criticizes the dominant focus on normative technical evaluation knowledge and competencies in the advancement of the evaluation professionalisation agenda and notes a lack of vigorous discussion of developing a professional ethos for evaluation. Professionalising evaluation as an occupation (and supply-and-demand commodity and service) with a focus on credentials and certification falls short of advancing evaluation as a public good of societal value. According to Schwandt (2017: 548), the professional ethos refers to the 'sum of a professional group's moral principles, core values, epistemic and aesthetic dispositions, and aspirations that each member of the group takes into consideration in interacting with others in a professional context'. Evaluation as a public work, as advanced by Schwandt and Reynolds: *"...combines insights [...] with ideas from critical system heuristics and the literature on knowledge utilization"* (Schwandt, 2017: 550; Reynolds and Schwandt, 2017).

1  
2  
3 If evaluation as a profession really wants to live up to the proclaimed goal of contributing  
4 worldwide to “*a transformed global community characterized by transparency, accountability,*  
5 *and progress towards the common good*” (EvalPartners, 2016:3), then, as Schwandt  
6 concludes, “*we need a much more public and energetic discussion of the professional ethos of*  
7 *evaluation, [...] what actually comprises its shared understanding of moral principles, values,*  
8 *aspirations, and ways of behaving*” (Schwandt 2017: 552).  
9  
10

11  
12 From the heritage of STiP, a focus on the evaluation practitioner in bringing about a second-  
13 order shift in evaluation practice can be one way to contribute to a professional ethos of  
14 evaluation practice. Second-order evaluation practice can be in support of enabling post-  
15 normal evaluation practice which is needed to respond to the post-normal characteristics of  
16 our time and the near future. Schwandt (2019) argues that the conventional understanding  
17 and practices of ‘normal’ evaluation are no longer sufficient. It is time for ‘post-normal’  
18 evaluation to come to the fore to more adequately deal with complex situations of change  
19 and uncertainty, he claims.  
20  
21

22  
23 Further professionalisation and capacity building for ethically reflexive, second-order systemic  
24 evaluation practice might also contribute to advance the still weak institutionalisation of  
25 systems thinking in organisations as well as rules and norms of conceptualising,  
26 commissioning, conducting and using evaluations. There are already opportunities in  
27 evaluation professionalisation and capacity building efforts which can be built upon e.g.:

- 28 • The Global Evaluation Agenda 2016-2020 (EvalPartners 2016) includes a chapter  
29 dedicated to the ‘strengthening of individual capacities for evaluation’ (Chapter A.3:  
30 21-28) including: ethical dispositions, professional autonomy, expertise and  
31 credentials (p. 21).<sup>4</sup>
- 32 • The innovative Voluntary Evaluator Peer Review (VEPR) pilot projects which have been  
33 implemented by the EES and UKES evaluation societies as a professional development  
34 service for their members. These are grounded in reflective practice principles, which  
35 as a central position in these professional societies’ evaluation capability frameworks  
36 lays a fertile foundation for further innovation/professionalization.  
37  
38  
39  
40  
41  
42  
43  
44

45 The next turn in this positive development may become the nurturing of second-order  
46 evaluation practice, by enhancing and expanding reflexive systemic evaluation i.e., second-  
47 order, systemic evaluation practice principles incorporated into capability frameworks. This  
48 requires extension of contemporary evaluation research agendas.  
49  
50  
51  
52  
53  
54

---

55 <sup>4</sup> The document also recognizes the important of ‘reflective practice’ for the individual evaluator’s capacity  
56 building. The evaluation community has also developed a number of evaluators’ capabilities frameworks  
57 focused on knowledge, practice and dispositions.  
58  
59  
60

### 4.3. An ongoing research agenda

There is an ongoing research imperative to effect the means and substance of a transformed, second-order, systemic evaluation praxis. The authors' own research agenda is directed at this imperative.

Research is currently under way to address and answer research questions that seek to elucidate the constraints and possibilities for transforming towards second-order practices by evaluation practitioners:

- How do evaluation practitioners engage with complex situations of change and uncertainty?
- How do evaluators reflect on the choices for approaches and methods in these situations?
- What opportunities exist for evaluators to make a 'second-order' shift in these situations?

The research process is designed as a set of nested learning systems grounded within and framed by situated practices. Using systemic action research principles, this research is firmly rooted in the experiential practice of the lead author (illustrated in vignettes 1 and 2), and is conducted through three modes, or cycles, of inquiry and with differing researcher positionalities: first-, second and third person inquiry (Torbert, 2001):

- the practitioner/researcher as 'evaluation and systems practitioner-conceptualizer' and researcher embedded within this ecosystem and praxis field (through continuous first-person inquiry ('learning for me');
- Practitioners engaged in the area of practice of evaluation (through second-person inquiry ('learning with others' as co-inquiry);
- Actors – and practices - in the wider situation of interest of developing (complexity-sensitive) public policy knowledge and systemic evaluation practice (third-person inquiry ('engaging in learning with a wider community').

This research is designed to 'walk the talk' of systemic practice by applying it to research and evaluation as a practice. The inquiry speaks to the broader developments and interests in the fields of public policy, science technology studies, and evaluation that are concerned with the question of how policies and decision-making can become more systemic in order to respond to accelerating complexity in a world that is now increasingly 'beyond the stable state' (Schön, 1971). Where traditional evaluation practices are becoming less effective in complex situations of change and uncertainty, evaluation practitioners must become more epistemologically and ontologically aware and better equipped to effect change that is systemically desirable as well as culturally feasible (Checkland and Scholes, 1990).

The proposed research is designed to avoid entrapment in a rigid systematic practice that assumes the linear transfer of knowledge. By moving towards third person action inquiry the



1  
2  
3 opportunity exists to use prior learning from these other modes as input into possible designs  
4 for learning systems enacted with other stakeholders i.e., knowing-in-action. In turn this  
5 creates the possibilities for emergent, contextualised, transformation instead of the  
6 mainstream focus on delivery and adoption i.e., knowledge transfer. Knowing-in-action can  
7 also be a means to enact that important systems concept – feedback – through participation  
8 in a learning system.  
9

10  
11 Learning from this research is expected of value for the theoretical advancement and  
12 contribution to knowledge of the evaluation discipline, but also of practical value to  
13 evaluators, to enhance their personal capacities to be, and engage, in complex situations of  
14 change and uncertainty, acting with systemically aware responsibility when doing evaluations.  
15  
16

## 17 18 19 20 **5. Conclusion: opportunities for second-order practice shifts in evaluation practice -** 21 **towards systemic evaluation practice and practitioners** 22

23  
24 Duffy (2017:149) argues that it “is only through remaining open to potential, yet unknown  
25 emergent transformations that the disciplinary and controlling effects of knowledge  
26 production processes can be unsettled”. We conclude that a second-order shift in evaluation  
27 praxis is not only justified, but necessary, as part of an unsettling project relevant to our  
28 human circumstances in the Anthropocene (Ison 2016).  
29

30  
31 A second-order shift helps to generate greater epistemological awareness in approaching  
32 evaluation and engaging evaluands as situations of interest that can be understood, utilised  
33 and transformed systemically. In contrast to traditional first-order evaluation practice, reality  
34 is brought forth relationally and experientially with participation and inclusion of others  
35 involved in a situation. ‘Evaluation is all about moving beyond where we are now’ (Duffy 2017:  
36 150).  
37  
38

39  
40 A second-order shift involves assuming at the outset that all situations / evaluands have  
41 elements of complicatedness, complexity and conflict. Key to such a shift is the role of the  
42 practitioner – the evaluator herself. In this understanding, evaluators or researchers are  
43 themselves part of the situation they seek to understand, change, evaluate, transform using  
44 systems as ways of knowing, inquiring or doing. A second-order shift brings awareness that an  
45 ‘as if’ position is possible – to see situations as systems to learn in a particular way. A non-  
46 reflexive commitment to first-order, systematic practice limits change possibilities and the  
47 focus of praxis to that of external observer of a system that is ‘out there’ (as with a mainstream  
48 external accountability perspective of evaluation).  
49  
50

51  
52 Evaluation practitioners are at the heart of systemic evaluation as understood in the tradition  
53 of STiP. Evaluators can develop their own praxis capacity and capability through reflexive use  
54 of the juggler isophor: their own Being, Engaging with the situations of interest,  
55 Contextualising their systemic evaluation performances and Managing their overall evaluation  
56  
57  
58  
59  
60



1  
2  
3 performance (BECM). Developing these capacities for systemic evaluation practice can be  
4 integrated into evaluation professionalisation and capacity building efforts, building further  
5 on reflective practitioner elements.  
6  
7

8 Future professionalism and the promise of efficacy of a second-order shift to systemic  
9 evaluation will need to deal with those systemic issues that plague the evaluation  
10 professional. Wadsworth (2010: 271) listed these as: (i) changes that are wanted but don't  
11 happen; (ii) changes that happen that people (stakeholders) don't want; (iii) decisions and  
12 politics that seem unresponsive, insensitive or prematurely pragmatic; (iv) inaccurate  
13 assumptions, over-generalisations, and an inability to see and hear what people are really  
14 saying; (v) preoccupation with fixing things that are going wrong with little or no time spent  
15 to make things right in the first place, or (vi) solutions becoming new problems ..'. To this we  
16 could add evaluations that merely tick a box or where the report sits on the shelf, where there  
17 is no feedback, learning and change. STiP-informed research and evaluation practice can  
18 support practitioners in knowing how to build supportive contexts for second-order systemic  
19 evaluation praxis –overcoming the constraints as well as institutionalising the enablers.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## References

- American Evaluation Society Topical Interest Group 'Systems in Evaluation TIG - A Topical Interest Group of the American Evaluation Association'. Available at: <https://www.systemsinevaluation.com/> (accessed 5 January 2020)
- American Evaluation Association (AEA). Available at: <https://www.eval.org/> (accessed 5 January 2020).
- Argyris C and Schön D (1974) *Theory in practice: Increasing professional effectiveness*. San Francisco: Jossey-Bass.
- Bamberger M, Vaessen J and Raimondo E (2016) *Dealing with complexity in development evaluation, a practical approach*. Thousand Oaks: Sage.
- Blackmore C, Foster N, Collins K (et al.) (2017) Understanding and developing communities of practice through diagramming, In Oreszczyn S and Lane A (eds) *Mapping environmental sustainability: reflecting on systemic practices for participatory research*: 155-182. Policy Press, London.
- Box G (1976) Science and Statistics. *Journal of the American Statistical Association*, 71 (356): 791-799
- Caffrey L and Munro E (2017) A systems approach to policy evaluation. *Evaluation*. London, England: SAGE Publications, 23(4): 463-478
- Cairney P (2016) *The Politics of Evidence-Based Policy Making*. London: Palgrave Macmillan Pivot.
- Catwell L and Sheikh A (2009) Evaluating eHealth interventions: the need for continuous systemic evaluation. *PLoS medicine*, 6(8): e1000126.
- Checkland P and Scholes J (1990) *Soft Systems Methodology in Action*. Chichester: Wiley.
- Derwisch S and Löwe P (2015) Systems Dynamics Modelling in industrial development evaluations. *IDS Bulletin* 46 (1): 44-57.
- Duffy DN (2017) *Evaluation and Governing in the 21st Century: disciplinary measures, transformative possibilities*. Palgrave Studies in Science, Knowledge and Policy. London: Palgrave Macmillan Pivot
- European Evaluation Society (EES). Available at: <http://europeanevaluation.org> (accessed 5 January 2020)
- European Evaluation Society (EES) Voluntary Evaluator Peer Review (VEPR). Available at: <http://europeanevaluation.org/community/thematic-working-groups/twg4/voluntary-evaluator-peer-review> (accessed 5 January 2020)
- European Evaluation Society (2018) 13th European Evaluation Society Biennial Conference, 'Evaluation for more resilient societies', Thessaloniki, Greece, 1-5 October 2018. Available at: <http://www.ees2018.eu/> (accessed 5 January 2020)

- 1  
2  
3 EvalPartners (2016) EvalAgenda 2020, Global Evaluation Agenda 2016-2020. Available at:  
4 <https://evalpartners.org/sites/default/files/files/Evalagenda2020.pdf> (accessed 5 January 2020)  
5  
6 Forss K, Marra M and Schwartz R, (eds) (2011) Evaluating the complex: attribution, contribution and  
7 Beyond. *Comparative Policy Evaluation* series (18). New Brunswick, N.J.: Transaction Publishers.  
8  
9 French RD (2018) Is it time to give up on evidence-based policy? Four answers. *Policy & Politics*, 47  
10 (1): 151-168  
11  
12 Funtowicz S and Ravetz J (1993) Science for the post-normal age. *Futures* 31: 735–55.  
13  
14 Gates E (2017) Towards Valuing with Critical Systems Heuristics. *American Journal of Evaluation*: 1-  
15 20  
16  
17 Hall JN, Ahn J and Greene JC (2012) ‘Values Engagement in Evaluation: Ideas, Illustrations, and  
18 Implications’, *American Journal of Evaluation*, 33(2): 195–207  
19  
20 IDS Bulletin (2014). 45 (6) November)  
21  
22 IDS Bulletin (2015). 46 (1) January)  
23  
24 Ison RL (2016) Governing in the Anthropocene: what future systems thinking in practice? *Systems*  
25 *Research & Behavioral Science* 33 (5): 595-613  
26  
27 Ison RL (2017) *Systems Practice: how to act, in situations of uncertainty and complexity in a climate-*  
28 *change world*. 2<sup>nd</sup> ed. London: Springer  
29  
30 Ison RL (2018) Governing the human-environment relationship: systemic practice. *Current Opinion in*  
31 *Environmental Sustainability* 33: 114-123. <https://doi.org/10.1016/j.cosust.2018.05.009> (accessed 30  
32 January 2020)  
33  
34 Ison RL and Blackmore C (2014) Designing and developing a reflexive learning system for managing  
35 systemic change. *Systems Education for a Sustainable Planet, Special Issue, Systems*, 2(2): 119-136  
36 (doi:10.3390/systems2020119).  
37  
38 Ison RL, Collins KB and Wallis PJ (2015) Institutionalising social learning: towards systemic and  
39 adaptive governance. *Environmental Science & Policy*, 53(PB): 105–117  
40  
41 Ison R and Russell, D (eds) *Agricultural Extension and Rural Development: Breaking out of Traditions*.  
42 Cambridge University Press, Cambridge, UK.  
43  
44 Ison R and Straw E (2020) *The hidden power of systems thinking: governance in a climate emergency*.  
45 London: Routledge.  
46  
47 Korzybski A (1933) *Science and sanity: an introduction to non-Aristotelian systems and general*  
48 *semantics*. Englewood, NJ: The International Non-Aristotelian Library Publishing Company.  
49  
50 Kusters C et al. (2019) Conference report: monitoring and evaluation for inclusive and sustainable  
51 food systems. 3-4 April 2019, Wageningen Centre for Development Innovation, The Netherlands.  
52 Wageningen University & Research. Report WCDI-19-066. Available at:  
53 <https://edepot.wur.nl/506604> (accessed 5 January 2020).  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Love AJ (1991) The process of internal evaluation. In: Love, AJ *Applied Social Research Methods: Internal evaluation*. Newbury Park, CA: SAGE Publications, pp. 36-62  
4  
5  
6 Maturana H and Poerksen B (2004) *From being to doing: the origins of the biology of cognition*.  
7 Heidelberg: Carl-Auer  
8  
9 Midgley G (2007) *Systems Thinking for Evaluation.*, In: Williams, B and Imam, I *Systems Concepts in Evaluation, an expert anthology*. Point Reyes, CA: American Evaluation Association, pp. 11-34.  
10  
11  
12 Mitchell A (2019) *Second-order learning in developmental evaluation: new methods for complex conditions*. Cham: Palgrave Pigot  
13  
14  
15 Ofir Z et al. (2019) From monitoring goals to systems-informed evaluation: insights from SDG14. IEED Briefing, March 2019. Available online: <https://pubs.iied.org/17706IIED/> (accessed 5 January 2020)  
16  
17  
18  
19 OpenLearn *Managing complexity: A systems approach – introduction*. Available at:  
20 [https://www.open.edu/openlearn/science-maths-technology/computing-and-ict/systems-](https://www.open.edu/openlearn/science-maths-technology/computing-and-ict/systems-computer/managing-complexity-systems-approach-introduction/content-section-0?active-tab=description-tab)  
21 [computer/managing-complexity-systems-approach-introduction/content-section-0?active-](https://www.open.edu/openlearn/science-maths-technology/computing-and-ict/systems-computer/managing-complexity-systems-approach-introduction/content-section-0?active-tab=description-tab)  
22 [tab=description-tab](https://www.open.edu/openlearn/science-maths-technology/computing-and-ict/systems-computer/managing-complexity-systems-approach-introduction/content-section-0?active-tab=description-tab) (accessed 5 January 2020)  
23  
24  
25 Open University (2010) *Study guide TU811 Thinking strategically: systems tools for managing change*.  
26 Milton Keynes: Open University.  
27  
28 Open University (2020) *MSc in Systems Thinking in Practice*. Available online:  
29 <http://www.openuniversity.edu/courses/postgraduate/qualifications/f47> (accessed 5 January 2020)  
30  
31  
32 Patton M (1994) Developmental Evaluation. *Evaluation Practice* 15 (3): 311-320  
33  
34 Patton M (2011) *Developmental Evaluation applying Complexity Concepts to enhance Innovation and Use*. New York, NY: Guilford Press.  
35  
36 Patton M (2018) *Principles-Focused Evaluation*. London/New York: Guilford Press.  
37  
38 Patton M (2019) Blue Marble Evaluation. Available at: [https://www.utilization-](https://www.utilization-focusedevaluation.org/blue-marble-evaluation)  
39 [focusedevaluation.org/blue-marble-evaluation](https://www.utilization-focusedevaluation.org/blue-marble-evaluation) (accessed 5 January 2020)  
40  
41  
42 Piirainen KA, Gonzalez RA, and Bragge J (2012) A systemic evaluation framework for futures  
43 research. *Futures*, 44(5): 464-474.  
44  
45 Ramalingam B (2013) *Aid on the edge of chaos. Rethinking international cooperation in a complex world*. Oxford: Oxford University Press.  
46  
47  
48 Reynolds M (2015) (Breaking) The iron triangle of evaluation. *IDS Bulletin* 46: 71–86.  
49  
50 Reynolds M et al. (2012) Complexity, systems thinking and evaluation - an emerging relationship?  
51 *Evaluation Connections, Newsletter of the European Evaluation Society*: 7–9.  
52  
53 Reynolds M et al. (2016) Towards Systemic Evaluation. *Systems Research and Behavioral Science* 33:  
54 662–673.  
55  
56 Reynolds M and Holwell S (eds.) (2020) *Systems approaches to managing change: a practical guide*.  
57 2<sup>nd</sup> ed. London: Springer.  
58  
59  
60

- 1  
2  
3 Reynolds M and Schwandt T (2017) Evaluation as public work: an ethos for professional evaluation  
4 praxis. In: *UK Evaluation Society Annual Conference: The Use and Usability of Evaluation:*  
5 *demonstrating and improving the usefulness of evaluation*, 10-11 May 2017, London, UK Evaluation  
6 Society.  
7  
8  
9 Rockefeller Philanthropy Advisors (2019) Scaling Solutions toward Shifting Systems Initiative:  
10 Assessing Systems Change: A Funders' Workshop Report July 29 - August 1, 2019. Available at:  
11 [https://www.rockpa.org/wp-content/uploads/2019/10/Assessing-Systems-Change-A-Funders-](https://www.rockpa.org/wp-content/uploads/2019/10/Assessing-Systems-Change-A-Funders-Workshop-Report-Rockefeller-Philanthropy-Advisors-August-2019.pdf)  
12 [Workshop-Report-Rockefeller-Philanthropy-Advisors-August-2019.pdf](https://www.rockpa.org/wp-content/uploads/2019/10/Assessing-Systems-Change-A-Funders-Workshop-Report-Rockefeller-Philanthropy-Advisors-August-2019.pdf) (accessed 5 January 2020)  
13  
14 Russell DB and Ison RL (2000) The research-development relationship in rural communities: an  
15 opportunity for contextual science. In Ison RL and Russell, DB (eds) *Agricultural Extension and Rural*  
16 *Development: Breaking out of Traditions*. Cambridge University Press, Cambridge, UK. pp. 10-31.  
17  
18 Schön D (1971) *Beyond the stable state: public and private learning in a changing society*. Random  
19 House.  
20  
21 Schön D (1984) *The reflective practitioner: how professionals think in action*. Aldershot, UK: Ashgate.  
22  
23 Schön D (1987) *Educating the Reflective Practitioner: towards a new design for teaching and learning*  
24 *in the professions*. San Francisco, CA: Jossey-Bass.  
25  
26 Schön D (2010) Government as a learning system. In Blackmore, C. (ed.) *Social Learning Systems and*  
27 *Communities of Practice*. Springer, Dordrecht, pp. 5-16.  
28  
29 Schwandt T (2015) Reconstructing professional ethics and responsibility: implications of critical  
30 systems thinking. *Evaluation*, 21 (4): 462–466.  
31  
32 Schwandt T (2017) Professionalization, Ethics, and Fidelity to an Evaluation Ethos. *American Journal*  
33 *of Evaluation*, 38(4): 546–553.  
34  
35 Schwandt T (2019) Post-normal evaluation? *Evaluation*, 25(3): 317–329.  
36  
37 Scriven M (1991) *Evaluation thesaurus*, 4th edition. Thousand Oaks, CA, Sage  
38  
39 Scriven M (1996) Types of evaluation and types of evaluator. *Evaluation Practice*. Elsevier B.V., 17(2):  
40 151–161  
41  
42 Scriven M (2001) Evaluation: future tense. *American Journal of Evaluation*, 22 (3): 301–307  
43  
44 Scriven M (2003) Evaluation Theory and Metatheory. In: Kellaghan T., Stufflebeam D.L. (eds)  
45 *International Handbook of Educational Evaluation*. Kluwer International Handbooks of Education, vol  
46 9. Springer, Dordrecht, pp. 15-30.  
47  
48 Stephens A, Lewis E and Reddy S (2018) Towards an inclusive systemic evaluation of the SDGs:  
49 gender, equality, environments and marginalised voices (GEMs). *Evaluation*, 24 (2): 220-236.  
50  
51 Stern E (2011) Editorial. *Evaluation*, 17(4): 323-325.  
52  
53 Torbert W (2001) 'The practice of action inquiry'. In P. Reason, Bradbury, H. (eds) *Handbook of Action*  
54 *Research, participative inquiry and practice*, pp. 250-260. London, SAGE Publications  
55  
56  
57  
58  
59  
60

1  
2  
3 UK Evaluation Society Voluntary Evaluator Peer Review (VEPR). (Online):  
4 <https://www.evaluation.org.uk/professional-development/voluntary-evaluator-peer-review/>  
5 (accessed 5 January 2020).  
6

7  
8 United Kingdom Government Office for Science (2007), Foresight: Tackling Obesities: Future Choices.  
9 Qualitative modelling of policy options. (Online):  
10 [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/295150/07-1183-obesity-modelling-policy.pdf)  
11 [295150/07-1183-obesity-modelling-policy.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/295150/07-1183-obesity-modelling-policy.pdf), and  
12 [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/296290/obesity-map-full-hi-res.pdf)  
13 [296290/obesity-map-full-hi-res.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/296290/obesity-map-full-hi-res.pdf) (accessed 5 January 2020).  
14

15 Ulrich W (1983) *Critical Heuristics of Social Planning: a new approach to practical philosophy*. Bern,  
16 Switzerland, and Stuttgart, Germany: Haupt. Paperback reprint version. Chichester, UK: Wiley, 1994  
17

18 Ulrich W (1996) *A Primer to Critical Systems Heuristics for Action Researchers*. Hull, UK, Univ of Hull,  
19 Centre for Systems Studies.  
20

21  
22 Ulrich W and Reynolds M (2020), '*Critical Systems Heuristics*', in Reynolds, M. and Holwell, S. (eds)  
23 (2020), *Systems Approaches to Managing Change: a practical guide*, London, Springer, The Open  
24 University.  
25

26 Vickers G (1970) *Value systems and social processes*. London, Penguin Books  
27

28 Vickers G (1978) *Responsibility – its sources and limits*. Systems Inquiry Series. Seaside, CA,  
29 Intersystems Publications.  
30

31 Von Förster H (1984) *Observing Systems*. Salinas, CA, Systems Publications  
32

33 Von Förster H (1992) Ethics and second-order cybernetics. *Cybernetics Human Knowing* 1:9-19  
34

35 Wadsworth Y (1997) *Everyday Evaluation on the Run*. 2<sup>nd</sup> edn. Sydney, Allen & Unwin.  
36

37 Wadsworth Y. (2010) *Building in Research and Evaluation. Human Inquiry for Living Systems*. Sydney,  
38 Allen&Unwin.  
39

40 Wadsworth Y (2016) *Everyday evaluation on the run: the user-friendly introductory guide to effective*  
41 *evaluation*. 3<sup>rd</sup> ed. Abingdon, Routledge.  
42

43  
44 Williams B (2013) Three core concepts: inter-relationships, perspectives, boundaries. *Evaluation*  
45 *Connections: Newsletter of the European Evaluation Society*, June 2013: 7-8.  
46

47 Williams B and Hummelbrunner R (2009) *Systems Concepts in Action, a practitioner's toolkit*,  
48 Stanford, Stanford Business Books.  
49

50 Williams B and Imam I (eds.) (2007) *Systems Concepts in Evaluation, an expert anthology*. Point  
51 Reyes, CA, American Evaluation Association.  
52  
53

**Abstract:**

The need for, and possibilities of, a second-order shift in evaluation practice are explored. Second-order evaluation practice enables an evaluator to improve practice as a skilled practitioner, acknowledging her embeddedness within an evaluand. The paper explores evaluation practice as experienced by professional evaluators, using ideas from developmental evaluation coupled with systemic evaluation in the tradition of systems thinking in practice (STiP). Systemic evaluation aims to capture systemic sensibilities – the bigger picture – of complex turbulent situations of change underpinning evaluands. Attributes of second-order practice with systemic evaluation are understood as being aligned with both systemic and systematic modes of evaluation praxis. Personal experiences are provided where this juxtaposing praxis has been found wanting. By example, a STiP framework is explored as heuristic support for making a second-order practice shift. The paper concludes with a discussion of some implications for developments in professionalising evaluation practice and research.

**Keywords** : developmental evaluation, evaluation praxis, evaluation professionalisation, isophor, second-order practice, systemic evaluation, systems thinking in practice.

**Résumé****Vers une évaluation systémique en période de turbulences - un changement de second ordre dans la pratique de l'évaluation**

La nécessité et les possibilités d'un changement de second ordre dans la pratique de l'évaluation sont explorées. Ce changement permet à l'évaluateur de donner un sens à sa pratique et de l'améliorer en permanence en tant que praticien compétent, reconnaissant sa propre intégration dans ce qu'il évalue. L'article rend compte d'un projet d'action-recherche actuellement en cours qui étudie la pratique de l'évaluation telle qu'elle est vécue par les évaluateurs professionnels, en utilisant les idées de l'évaluation développementale couplée à l'évaluation systémique dans la tradition de la Pensée Systémique mise en Pratique (PSP). L'évaluation systémique vise à saisir les sensibilités systémiques - la vue d'ensemble - des situations de changement complexes et turbulentes qui sous-tendent les sujets évalués. Les attributs de la pratique de second ordre avec l'évaluation systémique sont compris comme étant alignés avec les aspects à la fois systémiques et systématiques de la pratique de l'évaluation. Des expériences personnelles témoignent de situations où cette pratique juxtaposée s'est avérée insuffisante. A titre d'exemple, un cadre de PSP comprenant quatre dimensions - être, s'engager, contextualiser et gérer (EECG) - est exploré comme une source potentielle pouvant amener ce changement de pratique de second ordre. L'article conclut sur une discussion de certaines implications de cette recherche dans les contributions aux débats sur la professionnalisation de la pratique de l'évaluation.



**Titre:**

Vers une évaluation systémique en période de turbulences - un changement de second ordre dans la pratique de l'évaluation

**Résumé :**

La nécessité et les possibilités d'un changement de second ordre dans la pratique de l'évaluation sont explorées. Ce changement permet à l'évaluateur de donner un sens à sa pratique et de l'améliorer en permanence en tant que praticien compétent, reconnaissant sa propre intégration dans ce qu'il évalue. L'article rend compte d'un projet d'action-recherche actuellement en cours qui étudie la pratique de l'évaluation telle qu'elle est vécue par les évaluateurs professionnels, en utilisant les idées de l'évaluation développementale couplée à l'évaluation systémique dans la tradition de la Pensée Systémique mise en Pratique (PSP). L'évaluation systémique vise à saisir les sensibilités systémiques - la vue d'ensemble - des situations de changement complexes et turbulentes qui sous-tendent les sujets évalués. Les attributs de la pratique de second ordre avec l'évaluation systémique sont compris comme étant alignés avec les aspects à la fois systémiques et systématiques de la pratique de l'évaluation. Des expériences personnelles témoignent de situations où cette pratique juxtaposée s'est avérée insuffisante. A titre d'exemple, un cadre de PSP comprenant quatre dimensions - être, s'engager, contextualiser et gérer (EECG) - est exploré comme une source potentielle pouvant amener ce changement de pratique de second ordre. L'article conclut sur une discussion de certaines implications de cette recherche dans les contributions aux débats sur la professionnalisation de la pratique de l'évaluation.

**Mots-clés:** évaluation développementale, évaluation systémique, isophore, pensée systémique en pratique, pratique d'évaluation, pratique de second ordre, professionnalisation de l'évaluation.