

Dr. Deirdre Niamh Duffy is a senior lecturer in Interdisciplinary Studies at Manchester Metropolitan University

Scientism, governance and evaluation: Challenging the ‘good science’ of the UK evaluation agenda for youth work

Deirdre Niamh Duffy

Abstract

The UK government’s evaluation agenda for youth work has been subject of significant debate. Its opponents have highlighted the ends-oriented model of evaluation - Theory of Change - proposed by the UK government with youth work’s open-ended, fluid approach (Ord, 2014; In Defence of Youth Work, 2012). However, this model of evaluation has gained a great deal of traction among local authorities and youth work organisations (including the National Youth Agency). This is potentially due to the government’s claim that their aim is simply ‘good science’ using the language of ‘scientific rigour’ and ‘better assessment’. This article argues that the approaches being marketed – and adopted – are an example of regulatory ‘scientism’ wherein allusions to ‘good’ science are used to govern youth work. However evaluation of youth work should not be rejected entirely as it can also provide opportunities for innovation and change.

Keywords

Youth work, evaluation, policy, research

Introduction

Youth work and the daily practices of youth work are subject to increasing regulation. As numerous critics have noted, the practice is currently being subsumed within an overarching discourse of managerialism oriented towards achieving pre-set impacts and outcomes (Ord, 2014; Davies, 2013; Issitt and Spence, 2005). Connected to this, according to youth work commentators (Ord, 2014; Cooper, 2012), is the production of evidence through particular forms of evaluation research. Following the Department of Education and Skills (DfES) White Paper *Transforming Youth Work - Resourcing Excellent Youth Services* (DfES, 2002) and the House of Commons Education Committee’s (HCEC) review of youth work (2011), youth work organisations are now compelled to conduct ‘scientifically rigorous’ evaluations (ibid) which can clearly demonstrate the impact of their engagement with young people. Subsequent guidance has identified ‘Theory of Change’ evaluation as the preferred approach (Young Foundation, 2011).

This article is located within the debate about youth work evaluation in the UK as one strand of the broader discourse of neoliberal governmentality which the practice

operates within. However, while it is important to note, as others have (Issitt and Spence, 2005), that the current UK youth work evaluation agenda is a vehicle for governing youth work through research - a 'scientism' (Torrance, 2008: 522) or facade of systematic, scientifically rigorous analysis masking a decidedly non-scientific neoliberal political strategy - it is equally important not to reject the basic premise of this agenda, namely, that youth work should be subject to evaluation. While even the HCEC does not question the benefits participation in youth work brings to young people and communities, it would be wrong to assume that youth work is inherently more positive than other forms of engagement or that youth work is above critique. Youth work exists within and is the product of the same discourse as other organisations and thus embodies the same subjectivities and power relations. At the very least youth work's position as a more positive alternative to formal education (for example) need to be critiqued in order to gauge the validity of this claim. Moreover, a publicly-funded sector, youth work needs to articulate its merit and give an account of itself to the people who support it.

However, evaluation does not necessarily have to be restrictive in the way the current approach is. Arguably it is through such critical reflection that youth work can become aware of the power relations and disciplinary discourses within it and open up opportunities for innovation for new ways of thinking and being. As Foucault and Rabinow (1997) suggests it is through developing an awareness of ourselves that possibilities for 'mutation' or challenging discourse in its current form. Moreover, following Deleuze (1991) and the 'new empiricists', it is only through reflection that the abstract, imagined body can be brought into being in the present moment. In other words evaluation allows youth work to bring what it believes itself to be (its abstract self) to be known, touched, and embodied (its actual self). Further, in actualising the abstract (Bignall, 2007), youth work can bring itself closer to its future self (both known as an aspiration and unknown as a future aspiration). This self is addressed by Spinoza (Scruton, 1999) as the *not yet* or a self that exists in futurity if not in materiality.

What I contend here is that the approach being imposed upon youth work - Theory of Change - does not facilitate this and an alternative approach is needed. I do this by outlining the problems with the conflation of a single method with more rigorous scientific research (HCEC, 2011), demonstrating why Theory of Change is a poor fit for youth work and arguing that it facilitates 'governance through research' rather than productive evaluation contributing to 'social betterment' (Mark et al, 2000: vii). Throughout this discussion I adopt the labels 'scientism' or 'bad science' to describe the current evaluation model being imposed. This is in response to the claim by policy-makers and organisations like the Young Foundation that their structures produced more 'scientifically rigorous' evidence than the personalised accounts (labelled as 'anecdotal' by the committee) submitted during the HCEC review. That said, my aim here is not to present a unified model of 'good science'; as Kuhn (1953) notes the term 'science' covers a broad *disunified* range of approaches and perspective. Rather my aim is to suggest – albeit tentatively – a way to improve youth work evaluation which does not replicate the regulatory impacts of the current *scientistic* approach without rejecting evaluation entirely. I will try to open up a discussion on how to create a

dialogue between evaluation and policy so that the critical possibilities of evaluation can be exercised while still being useful to policy-makers.

Centrally the position this argument takes is twofold. First that the notion of 'good science' alluded to by the UK government under their evaluation agenda for youth work is principally another mechanism for reconfiguring youth work towards a more 'manageable' simple system oriented towards meeting pre-set goals. This is demonstrated by the poor fit of the model promoted - Theory of Change. Second the promotion of Theory of Change is an example of methodological fundamentalism - where the method is more important than the subject of evaluation. The article highlights how such fundamentalism is problematic not only because it subverts the research process - wherein methods are selected based on their appropriateness - but also because, following thinkers such as Oakley (2000), it silences particular types of evidence and experiential knowledge. Furthermore, following Foucault and Deleuze, it limits the ability of evaluation to facilitate critical thought that could potentially lead to innovation in youth work practice.

Theory of Change and youth work

Following the demand by the HCEC review for better evaluation of youth work, a framework document – *A framework for positive outcomes for young people* (Young Foundation, 2012) – was published. Commissioned by the Department of Education and produced by the Young Foundation, with the support of a consortium of youth work-oriented agencies, local authorities and private sector firms – this document outlines a model for conducting 'scientifically rigorous' evaluation of youth work's merit. The framework focuses on social and emotional capabilities organised in a seven point cluster and sets out 'a matrix of available tools for measuring these capabilities [and] outlines a step by step approach in practice' (Young Foundation, 2012: 4-5). Central to this evaluation matrix is the identification of desired outcomes and outputs by programme managers at the beginning of the intervention. The rationale for this is threefold. According to the Young Foundation report, *a priori* identification of outcomes and outputs will:

- *Clarify what the programme is trying to achieve (content) and how (process)*
- *Establish where the programme is working well and where further improvements are needed*
- *Close the loop with feedback on progress against business needs* (Young Foundation, 2012: 24)

The framing of 'scientifically rigorous' evaluation as involving the pre-identification of outcomes and outputs is further entrenched by the Cabinet Office's flagship programme for improving evaluation in youth work: the Centre for Youth Impact (CYI). Launched in November 2014 and proposed as a support agency for youth work professionals looking to develop their research skills, the CYI will 'provide overarching support for all impact measurement initiatives that are relevant to the youth sector' (Cabinet Office, 2014). Importantly for this article, despite claiming that the methodological frameworks they suggest to professionals will be 'bespoke', the first stage to conducting a good

evaluation according to Project Oracle, one of the three key organisations leading the CYI, involves identifying desired goals and outcomes. This is made explicit on Project Oracle's own advisory website where they explain evaluation as requiring the clear identification of outcomes (Project Oracle website). Project Oracle argues that the 'validity' of an evaluation can be assessed by comparing the evaluation against what they call the Standards of Evidence (Project Oracle website). An evaluation that meets the minimum standards of scientific rigour, they argue:

has provided a coherent and plausible description of the logic that lies behind it. This includes a description of the project activities, intended outcomes and aim, how these are connected and what assumptions are being made (Project Oracle, 2014)

Here again 'scientifically rigorous' evaluation is constituted as research which both identifies desired outcomes and outputs in advance and illustrates whether these outcomes and outputs were successfully achieved.

Methodologically this follows a Theory of Change model of evaluation research (something which Project Oracle make explicit on their website). First designed by Carol Weiss of the Aspen Institute (Weiss, 1995; Connell et al, 1995), Theory of Change (ToC) is a framework for conducting evaluation of social programmes which encourages programme stakeholders, in collaboration with evaluators, to make explicit the 'long-term vision of the initiative' and 'consider the *necessary* outcomes that will be required by the end of the programme if such an aim is to be met' (Blamey and Mackenzie, 2007: 443). Importantly ToC is not intended to act solely as a tool for social inquiry, ensuring accountability or assessing value/merit – three core features of evaluation according to researchers like Greene et al (2004), Alkin (2004), Feinstein (2012) and others – but as a means of assisting organisations to produce and put into practice evaluation research. ToC ostensibly makes evaluation practicable by constructing a feedback loop within the programme. In addition to encouraging stakeholders to make the long-term vision and necessary outcomes explicit, a ToC approach demands that 'those involved in the programme consider the most appropriate activities or interventions required to bring about the desired change' (Blamey and Mackenzie, 2007: 443). In doing so ToC effectively creates a roadmap for programme stakeholders to follow and refer back to as the programme is implemented.

Advocates of ToC (Weiss, 1995; Connell and Kubisch, 1998) argue that this approach to evaluation can 'sharpen the planning and implementation of an initiative' and 'helps guide choices about how and when to measure [activities' impact]' (Connell and Kubisch, 1998: 2). ToC is also framed as useful in providing 'actionable' evidence, i.e. evidence which can be 'deemed adequate and appropriate for guiding actions in real world contexts' (Julnes and Rog, 2009: 96). This is due to the fact that the process of designing a ToC evaluation involves constructing both a programme and an implementation theory (Weiss, 1995) which stakeholders can follow to achieve the pre-stated outcomes and aims.

There are a number of problems with the promotion of ToC evaluation in the context of youth work. Some of these problems have been alluded to by youth work advocates – the work of In Defence of Youth Work campaign, Cooper, Issitt and Spence, and Ord are particularly notable here. What I want to interject into this body of writing is that the central problem with applying ToC to youth work is that it reconstructs a complex system into a simple one, thereby serving to *regulate* youth work to meet particular targets rather than *evaluate* the merit of youth work more generally.

While what constitutes a ‘simple’ system is still the subject of much debate within evaluation science (Mowles, 2014), there has been a degree of consensus over what the core features of a simple system within the context of social policy are. These can be labelled as regularity, linearity and movement towards equilibrium (Callaghan, 2008). A simple system evolves at a steady rate along a continuous path towards a point of equilibrium. Coupled with this regularity, a simple system is characterised by its ‘boundedness’. A classic example used by systems theorists is that of a drug treatment regime for a specific health complaint. A drug treatment regime is designed with a fixed end point or goal in mind. The desired results and point at which the system can be considered complete are planned in advance.

ToC approaches, while presenting themselves as useful tools for evaluating complex social programmes (Connell and Kubisch, 1998) implicitly construct the programmes they investigate as simple systems. This is demonstrated by the fact that both identifying desired outcomes and designing an implementation and programme theory pathway or roadmap which will lead to the achievement of those outcomes are key stages in conducting a ToC evaluation (Weiss, 1995; Connell and Kubisch, 1998). The guidance and statements published by the UK government on the topic of youth work evaluation adopt and advocate the wider adoption of this methodological approach. The Framework for Outcomes, for example, outlines a model for youth work organisations to follow – a programme theory – in order to improve young people’s emotional and social capabilities. Conceptualising youth work as a simple system in this way suggests that, through good practice, young people can be impacted by interaction with youth work organisations and professionals in a consistent fashion and that these interactions, when of a consistent standard, can result in the achievement of specific positive outcomes.

Despite the range of outcomes the Framework identifies, the system it constructs is a relatively simple one. It suggests the possibility for a linear, causal relationship between the implementation of particular practices and the achievement of outcomes to exist. This argument is further entrenched by the Framework which depicts a model of working with young people which can, through implementing particular practices, result in specific outcomes (the ‘seven clusters’ of emotional and social capability) (Young Foundation, 2012). This is a fundamental misunderstanding of youth work which according to authors such as Young (1999) and Sercombe (2010) is much more fluid and responsive - taking young people as its primary constituent (Sercombe, 2010) and designing approaches and activities around their needs. It therefore represents a complex rather than a simple system.

Complex systems are very different to their simple counterparts. According to commentators such as Callaghan (2008) and Gilchrist (2000), complex systems are characterised by a lack of regularity, linearity and movement towards equilibrium. They operate, in the words of Barnes et al (2003), 'at the critical point on the continuum between rigidity and randomness characterised by a state of 'untidy creativity' (Barnes et al, 2003: 278). Unlike simple systems, such as that depicted in the Framework for Outcomes, there is no underlying roadmap which complex systems align themselves to. On the contrary, complex systems are constantly moving towards 'the edge of chaos' (Gilchrist, 2000: 267). Importantly, the identification of systems as complex by theorists such as Gilchrist is not intended to be a process of clarification or labelling. Gilchrist and others, in outlining why differentiates complex from simple systems also advocate from greater acceptance of complexity approaches. The 'untidy creativity' and 'chaos' which characterises complex systems, according to complexity advocates, enables periods of 'exponential' development (Callaghan, 2008), encourages innovation and openness to new ways of thinking and acting (Gilchrist, 2000).

While there is still a good deal of discussion on what characterises a complex system, some attempts to identify their chief characteristics has been made. As a general guide Cilliers (2005) argues that complex systems are dynamic, unbounded and non-linear. They are self-referential in the sense that order is negotiated as the system evolves. Their functioning is not reducible to their parts – there is no direct link between the components of the system and its operations. Finally, there are no causal pathways within complex systems. Any developments which emerge are erratic, unpredictable and cannot be explained as resulting from specific actions. Complex systems - although undesirable to proponents of approaches such as the Framework for Outcomes which campaigns for the production and achievement of regular, pre-set goals – are increasingly common within social programmes. This is due the fact that social programmes are now designed to operate at multiple levels, affecting for example individuals, communities, national populations and international populations (Barnes et al, 2004). Moreover, social programmes are intended to be adaptive to emergent issues and changes in conditions and to be open to avenues for programme development which open unexpectedly and opportunistically (Benjamin and Greene, 2009).

The model of complex systems outlined above has clear resonances with youth work practice. Youth work, as Sercombe (2010) argues, does not follow a particular framework but is always in a state of emergence. Youth work organisations' practices are constantly being reconstituted in collaboration with the young people they have contact with (Sercombe, 2010; In Defence of Youth Work, 2011). Youth work, according to its advocates and commentators, is an 'iterative' (Ord, 2014) practice which eschews forward-planning or fixed targets, arguing that these inspire instrumentalism (ibid) and reproduce the coercive relations of power between practitioners and young people that youth work is attempting to subvert (Sercombe, 2010). Moreover, within youth work, the scale of outcomes or impact is not reducible to the type of activity or approach adopted. The intensity of and effort needed to sustain a particular invention may not be reflected in the outcome of the intervention. Seemingly small practices – a game of pool, a

conversation – can yield benefits for the young person which far exceeds the resources supporting the activity (although there is of course a skill to mobilising these practices effectively). Given the disavowal of specific targets and opportunistic ways of working adopted by youth work, the application of ToC-resonant frameworks is a mistake.

Governance, 'good science' and the UK evaluation agenda

The above discussion indicates the dissonance between youth work and ToC evaluation, highlighting how the latter restructures youth work artificially into a simple format which is at some distance from its established complex form. However, this article is not just intended to problematise a specific method but challenge the evaluation agenda more generally through demonstrating that this agenda uses the notion of 'good/bad science' as a means to govern youth work rather than either develop further understanding of it or facilitate the emergence of better approaches.

Centrally what I want to demonstrate (and criticise) is that the current evaluation agenda for UK youth work is overly concerned with the application of a specific method - Theory of Change - and ignores the pluralist nature of social inquiry. Resultantly, youth work evaluation does not generate concepts and data which can further our understanding of the nuances of youth work and limits the capacity of evaluation to assist in the emergence of innovative ways of working or developing new knowledges. Crucially, this is intentional. It is another vehicle for a neo-liberal inspired programme for youth work targeted at structuring and orientating the practice in a particular direction (one which will yield pre-set outcomes). 'Science' within this discourse is there to legitimise and reinforce policy and assist in regulation. This use of 'science' to regulate complex systems is not novel; Oakley (2000), Denzin and Giardina (2006, 2010) and Trinder and Reynolds (2000) all describe how campaigns to introduce more 'scientifically rigorous' research in health and social care and education are mobilised as mechanisms for legitimising particular policy approaches - e.g. No Child Left Behind in the United States - and regulate interactions between practitioners and service-users at a local level.

However, this is not to say that all evaluation is regulatory or that youth work should not be requested to provide an account of itself. It would be a mistake to argue that we should simply accept youth workers' promotion of their practice as positive without requesting supporting evidence from them. This is not my intention. Indeed, as I will note later on, evaluation can be a positive experience as it encourages organisations to critically reflect on their practices drawing out, as Deleuze (1991) notes, the points of disjuncture and tensions. Following Foucault (Foucault and Rabinow, 1997) it is at these points that possibilities for new ways of interacting with young people can emerge and problematic relations between subjects (youth worker - young person, young person - adult and so on) can be disrupted. Furthermore, given that it is often publicly funded, youth work has a responsibility to demonstrate how the financial resources it draws upon are used. That said, as before, what I want to demonstrate is the problems with the current evaluation agenda.

To initiate this discussion I want to consider three points: (i) why the promotion of one method over others has been (and should be) opposed; (ii) how such methodological fundamentalism ignores potentially beneficial pluralist forms of evidence and inquiry, and (iii) how it restricts the critical possibilities presented by evaluation research and serves to regulate organisations. To assist in this latter point I will draw from broader writing on 'new empiricism' which explore research of 'sense-data' or experience more generally, positioning evaluation research as part of this.

Accusations and discussions of 'methodological fundamentalism' have been most prevalent in writing on the application of Randomized Control Trials (RCTs) in social programme evaluation (particularly educational programmes). Drawn from the natural sciences RCTs' objective is to demonstrate and identify the outcomes and impacts of particular programmes - and their socially curative potential - through the application of treatments in structured experiments (Torrance, 2010; Donaldson et al, 2009; Oakley, 2000). As a number of authors from the US and UK have argued, from the early 2000s on, RCTs have been increasingly positioned as the method *de rigueur* in evaluation research. While their promotion has been more explicit in the US (Shaker and Ruitenberg, 2007), analyses of guidance on how to ensure evaluation research quality in evaluation research point to similar trends in the UK (Torrance, 2010) albeit through the language of 'pilot programmes' (Cabinet Office, 2003) and command that 'findings/conclusions [need to be] supported by evidence' (Cabinet Office, 2003: 22).

The complaint of critics of RCTs is not that they are an inherently 'bad' social science research method or incompatible with social programme evaluation. Indeed Howe (2009) explicitly warns against an opposition of RCTs on these grounds as risking the construction of an equally problematic dogma of 'scientific' versus humanities-inspired approaches. Rather the central criticism is that the 'heavy promotion' of a particular methodology over all others:

... [reverses] the logical and time-honoured order of decisions when conducting social inquiry. First, one identifies the inquiry purposes and questions, and only then selects a methodology that fits these purposes and questions (Greene, 2009: 157; see also Chelimsky, 2007)

By emphasising a specific methodology, policy-makers marginalise questions of relevancy or desirability - whether the method is the most useful tool for assessing the value of particular approaches or whether the information gathered will be what policy-makers want (Torrance, 2010). As Shaker and Ruitenberg (2007) write:

It is quite possible to design random experiments [...] in the field of education - but nothing in that test design itself guarantees the desirability or relevance of the research questions or outcomes (Shaker and Ruitenberg, 2007: 211)

Fundamentally, opposition to RCTs is rooted in a professional ethic which resists the subservience of research to policy, i.e. that research should deliver what policy-makers want, arguing that the maintaining quality in social science research 'cannot be

ensconced in a single research method' (Torrance, 2010: 73). The pre-eminence of RCTs as a 'gold standards', authors such as Denzin and Giardina (2006) argue, presents 'scienticism' as science.

Importantly the promotion of RCTs - or indeed any single outcome-oriented method - is not just problematic in terms of preserving the norms of the social science research process, it also has a profound impact on the subjects of research, serving to discipline them to masculinist logics of 'hard' 'reliability' (in terms of their regularity in producing particular outcomes) as opposed to 'softer', more feminine, 'unreliability' (Oakley, 2000). The foci of evaluations are turned into objects which can be controlled and manipulated in order to yield outcomes which can be measured and compared. As Oakley's (2000) feminist critique of gender and method in social science research argue 'experiments are inherently 'objectifying' and alienating' (Oakley, 2000: 38) and lead to 'all bodies, whether human or animal, being seen as machines' (ibid: 39).

Connected with this methodological fundamentalism, the government's evaluation agenda for youth work is problematic due to the lack of recognition it gives to the pluralist nature of social scientific inquiry and indeed the 'disunity' of science (Kuhn, 1953). There are any number of methodological approaches which can be used to gather evidence of a social phenomenon. Despite the so-called 'paradigm wars' (Taylor and Balloch, 2005: 1) between advocates of different approaches, there is a general consensus across contemporary philosophies of science and social inquiry that there is no 'master epistemology' (St. Pierre, 2002) which can be universally applied. The general acceptance that the idea of a perfect source of evidence is illogical is reflected not least in the promotion of 'critical thinking' amongst the scientific community and need for researchers to be explicit about the faults in their own research designs and approaches.

On one level the UK government appears to recognise the pluralist nature of social science. Both the HCEC report and the Framework for Positive Outcomes criticise the supposed over-reliance on quantitative measures of participation – 'headcounts' in particular – in evaluation of youth work and advocate the inclusion of more diverse forms of evidence. At the same time, this seeming commitment to pluralism is accompanied by the labelling of evidence submitted to the HCEC review as 'non-scientific' or 'anecdotal' (HCEC, 2011) and the promotion of a very specific metric for measuring youth work organisations' impact through the Framework. Ironically, in rejecting oral and written submissions by young people and youth workers which illustrate the transformative effect of youth work through case studies and storytelling as insufficiently 'scientifically rigorous', the government evaluation agenda promotes the notion of a 'master epistemology' a belief which, according to contemporary writing on pluralist social inquiry, is indicative of scientific reductionism (Greene, 2009; St. Pierre, 2002).

The exclusion of submissions to the HCEC from the umbrella of scientifically rigorous evidence is underpinned by an implicit prioritisation of particular ways of knowing and forms of knowledge over others without much justification as to why. Such prioritisation

appears all the more confusing given that, in the Framework for Outcomes and description of the CYI, the government promotes gathering evidence from stakeholders. However, there are two potential explanations for the rejection of stakeholder evidence by the HCEC. The delineation between evidence from stakeholders submitted voluntarily to the HCEC as 'unscientific' and evidence gathered through the Framework for Outcomes matrix as 'scientific' reflects a specific understanding of the dynamics of knowledge production whereby knowledge is not constructed by participants but gathered from them. Such an understanding follows a discourse of credible knowledge and scientific inquiry which prioritises the figure of the external, value-neutral 'expert' as the only legitimate collator of information (Donaldson et al, 2009). This is much discussed and criticised by thinkers such as Oakley (2000) who argues that it reflects deeply entrenched social biases and unequal power relations between communities who experience a particular phenomenon and supposedly more learned scientists.

The treatment of knowledge gathered as of greater scientific value than knowledge volunteered can also be found elsewhere in UK policy. The Green Book: Appraisal and Evaluation in Central Government (HM Treasury, 2011), for example, states that participatory research methods where knowledge is co-constructed with participants rather than gathered from them through researcher designed matrices are not legitimate approaches in evaluation research. This is despite the extensive literature highlighting not just the methodological appropriateness but scientific legitimacy of participatory research methods (Weis and Fine, 2004; Greene et al, 2004; Whyte, 1991) for developing better understandings of the dynamics of social programmes. Moreover, advocates of participatory methods have consistently argued that these methodologies are more 'actionable' and more useful for facilitating social programme development and change than traditional researcher-directed data collection approaches. By depicting the HCEC submissions as 'unscientific' the UK government are, at best, inaccurately representing accepted norms of good practice in social scientific research which do not discount evidence on the basis of whether it was gathered or volunteered.

Another, interconnected, explanation for the disavowal of submissions to the HCEC as 'anecdotal' rather than 'scientific' relates to the politics of evaluation (Taylor and Balloch, 2005; Greene, 2009). This explanation has less to do with how knowledge is understood and more to do with the politicisation of research. According to Greene (2009), within contemporary policy debates, evidence is labelled as not meeting the standards of 'scientific rigour' when it does not facilitate the advancement of particular political projects, not when it is representative of 'bad science'. In the context of youth work and evaluation, this argument has already been outlined in some depth by In Defence of Youth Work (2011), Ord (2014) and others. These commentators argue that the HCEC rejected submissions indicating the 'transformative' effect youth work participation has as they did not sit well with contemporary neoliberal-inspired discourses of public and social policy management which insist on organisations and services becoming oriented towards the achievement of targets (In Defence of Youth Work, 2011). From this perspective the standards of 'scientific quality' applied by the UK government were directed not by contemporary philosophy of science and social inquiry but by political ideology.

The third issue which I wish to consider in this critique of the blanket promotion of ToC is the critical possibilities empirical research presents and how the current agenda undermines these. Here I will draw on debates regarding 'new empiricism'. Rooted in the writing of Gilles Deleuze - particularly his commentary on the work of David Hume (Deleuze, 1991) - this emerging area of discussion considers (amongst other things) the role of theory in empirical research. According to Deleuze, a key function of empirical inquiry is to connect disparate, partly-unknowable bodies which exist in abstraction with the concrete world and in doing so highlight the disjunctures between the body in abstraction and the 'actualised' body. Importantly, new empiricism and Deleuzian theory does not limit the term 'body' to 'a particular discrete entity [...] but to any form, including both material bodies and bodies of knowledge or ideas' (Bignall, 2007: 202) and as such can be applied to studies of organisations or practices such as youth work.

Taken into the context of evaluating youth work, a new empiricist perspective positions research not as a way of producing a verifiable generalised theory of how youth work organisations operate but a mechanism for generating new forms youth work can take or new practices it can engage in. The purpose of evaluation - even in its summative form - is not to surmise what youth work is but provide a means of articulating youth work as a complex array of meanings and positionalities. As Bignall (2007) explains, new empiricism argues that through empirical reflection:

the sites in one's own self where one's identity is multiple and perhaps contradictory [...]. These apparent points of disjunction in one's own identity [...] signal point where the constituting discourses are unstable. It is in seeking out and finding such points of instability, Deleuze and Guattari suggest, that a person becomes most advantageously placed to experiment with the assemblage they embody and with the [abstract body] they might access in order to become identified otherwise (Bignall, 2007: 210)

Concepts assist in this process by acting as 'thought experiments' (Adkins and Lury, 2009: 12) which help think through the conditions necessary for abstract bodies to come into existence and recognise the points of 'undecideability' (Bignall, 2007: 210) within them. They should not be complete representations of how organisations work or direct what organisations should do but facilitate critical thinking about the zones of undecideability within an organisation where new practices or new ways of being could be explored and experimented with.

However, the promotion of ToC undermines this extension and experimentation. Under this method reflection is not used as a mechanism for generating mutation or critical openness (Foucault and Rabinow, 1997) but as a corrective (Deleuze, 1991) - the aim is not to open spaces for experimentation but to ensure adherence to pre-set understandings. The manipulation of 'scientism' in order to restrict the capacity of organisations to experiment and introduce new ways of working is exemplary of how organisations are 'rendered governable' in contemporary society (Carmel and Harlock, 2008). By consistently positioning ToC/outcomes-oriented evaluation as of a higher

standard than participatory approaches (through labelling the results of the latter ‘anecdotal’ and ‘non-systematic’), government agencies have constructed a discourse of evaluation research where openings for stakeholders in youth work to express their experiences or challenge the targets set by government are closed down.

In adopting the language of ‘science’, the political nature of the UK evaluation agenda – highlighted in great depth by youth work commentators (Issitt and Spence, 2005; Ord, 2014) – is both being played down by government agencies and accepted by local authorities. That said, as the In Defence of Youth Work campaign indicates, the evaluation agenda is being rejected and questioned by some youth work organisations. However the list of signatories to Framework for Outcomes demonstrates that the model of ‘science’ peddled by government is also finding a good deal of support among national youth work agencies.

A way forward?

In light of the preceding arguments - particularly the implication that the methodological fundamentalism demonstrated by youth work evaluation policy is a strategy of regulation - it would be easy to resist any form of evaluation. However, such a knee-jerk reaction is equally problematic as the policing through campaigns for scientific rigour in that it ignores the very real pressure on governments to demonstrate that publicly financed programmes are productive (Torrance, 2010). It also ignores the potential of evaluation to assist campaigns for social justice and democratic dialogue (DiStefano, 2000). Youth work scholars have already recognised the potential for research practice to encourage young people’s critical learning and political participation (Coburn, 2011; Gormally and Coburn, 2014) and for evaluation specifically to strengthen the socially transformative possibilities youth and community organisations present (Cooper, 2012). If the problem then is the form of evaluation being projected as ‘good science’ not the aim of evaluating youth work per se then it is crucial to explore alternative research practices which do not artificially structure youth work into a simple system or act as a ‘corrective’. This is a daunting task and one which requires a much longer conversation. But in reaction to the arguments introduced above I wish to make two suggestions for how to circumvent the disciplinary effects of the current discourse of scientifically rigorous evaluation espoused in the UK’s agenda for improving youth work.

Based on the critiques of methodological fundamentalism and ‘gold standards’ (as silencing and exclusionary) presented by Oakley (2000) and others, the starting point for a discussion on how to integrate evaluation with youth work without restricting the latter should be to unsettle evidence as the main concern. RCTs and ToC are manifestations of a discourse of science obsessed with the application of methods in order to yield particular forms of evidence. Indeed the HCEC’s report and the CYI are both dedicated to the production of evidence. This is decidedly wrong-headed given the relativistic and malleable nature of data (a discussion covered in some depth by MacLure (2009) in her analysis of ‘the ruins’ of evidence), the dependence of evidence on interpretation and the inherent incompleteness of evidence (Deleuze, 1991).

As an alternative to this, a key focus should be developing an understanding of the different elements of youth work – an exercise of shining a light on or ‘stripping back non-cognition and sliding towards cognition’ (Gregg and Seigworth, 2013: 14) – while at the same time accepting that not everything will be ‘knowable’. The aim is not to gather particular forms of evidence but to come close to knowing as much about youth work as possible. This will not just involve broadening acceptance of qualitative research such as storytelling (an approach used by the In Defence of Youth Work campaign) but also lead to the forms of quantitative data-gathering which the HCEC dismisses as ‘counting’ being valued for the demographic information they provide. As Oakley (2000) notes, numerical evidence is not in itself problematic so long as it is not promoted over other forms of evidence in an artificial dualism of ‘good’/‘bad’ science.

The second tentative step towards a less disciplinary evaluation agenda would be to unsettle measuring economic ‘effectiveness’ in favour of fiscal transparency. These are conflated within the UK government’s evaluation guidance as part of the ‘three e’s’ – economic, effective, efficient – but as a focal point for evaluation research have very different impacts on the subject of evaluations. Fiscal transparency is in essence an exercise in description (what was money spent on, by whom, to what end) whereas economic effectiveness measurement is an exercise in monitoring returns for investments. Given its complex nature and the range of variables impacting upon its outcomes (i.e. the neighbourhood it operates in, the type of young people involved) youth work’s ability to clearly demonstrate economic effectiveness according to a cost-benefit analysis is limited. However economic effectiveness is not the only indicator that public expenditure is managed well. Transparency in what spending takes place can also give a clear indication of the work done in and by youth work organisations and allow policy-makers to meet the basis expectation that the public will be able to know the destination of their investments.

The fundamental difference between economic effectiveness and fiscal transparency relates to *determinism*. An evaluation dedicated to measuring economic effectiveness is focussed at determining whether spending was warranted or well-managed; an evaluation focussed at fiscal transparency is more interested in uncovering and describing what went on rather than assessing whether money could have been better spent directly (although this may well be one of the ways such an evaluation is subsequently used). Less disciplinary evaluations, according to theorists such as Mark et al (2000) should eschew determination:

Determination implies a role for evaluation that cannot be justified. It suggests coming to a fixed answer or settling a question. The very origins of the word suggest coming to an end. In evaluation, these connotations are unfortunate and undesirable. For one thing the limits of evaluation argue against using a term that implies finality (Mark et al, 2000: 55)

This connects back to the arguments made above, that the ToC model being imposed on youth work is problematic as it promotes a focus on ends – or finality – rather than innovation or progression.

Conclusion

As this article demonstrates the model of 'good science' proposed under the UK government's evaluation agenda for youth work is highly problematic. It artificially constructs an image of youth work as a simple system, overlooking its complexity. As a result, rather than allow organisations to articulate and shine a light on the complex dynamics of their work, as others have noted, it forces youth workers to focus on achieving pre-set goals. This runs counter to youth work's professional ideology of responsiveness and non-coercive relationships (Sercombe, 2010). Moreover, in addition to this managerialism, the evaluation agenda is overly reliant on a specific method – ToC – and limits the potential for evaluation to facilitate innovation and critical thinking. It is a methodological fundamentalist 'corrective' approach. To illustrate the problems with single-method reliance the article has looked to critiques of Randomised Control Trials as a 'gold standard' in educational research, drawing on the work of Oakley (2000), Torrance (2008, 2010) and others, arguing that this is a form of 'scientism' where the application of particular methods takes precedence over the subject of research on the grounds that 'science is science is science' (Lather, 2004).

However, the article has also recognised that, as youth work is publically funded, there is a political imperative for policy-makers to evaluate the practice. For that reason it has not discounted evaluation but proposed two ways that the current problematic discourse of 'good science' could be disrupted: displacing evidence as the main concern and replacing economic effectiveness with fiscal transparency as an objective of evaluation research. These tentative suggestions are intended to flow into a larger, on-going conversation about how to find a 'nexus' (Gormally and Coburn, 2014) between research practice and youth which may lead to transformation (Cooper, 2012).

References

- Adkins, L., and C. Lury (2009) "Introduction: What Is The Empirical?", *European Journal of Social Theory*, Vol. 12, No. 1, pp. 5–20
- Alkin, M. (ed.) (2004) *Evaluation roots: Tracing theorists' views and influences*, Thousand Oaks, CA: Sage Publications
- Barnes, M., Matka, E. and Sullivan, H. (2003) "Evidence, Understanding and Complexity: Evaluation in Non-Linear Systems", *Evaluation*, Vol. 9, No.3, pp. 265-284
- Benjamin, L.K. and Greene, J.C. (2009) "From Program to Network: The Evaluator's Role in Today's Public Problem-Solving Environment", *American Journal of Evaluation*, Vol. 30, No. 3, pp. 296-309
- Bignall, S. (2007) "A superior empiricism: The subject and experimentation." In *Pli*, vol. 18, pp. 201-217
- Blamey, A. and Mackenzie, M. (2007) "Theories of Change and Realistic Evaluation: Peas in a Pod or Apples and Oranges?", *Evaluation*, Vol. 13, No. 4, pp. 439-455
- Cabinet Office (2003) *Quality in qualitative evaluation. A framework for assessing research evidence*. London: Government Chief Social Researcher's Office
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/49

- [8322/a_quality_framework_tcm6-38740.pdf](#) (accessed: 08/08/2016)
- Cabinet Office (2014) "The Centre for Youth Impact Launches", <https://www.gov.uk/government/news/the-centre-for-youth-impact-launches> (accessed: 08/08/2016)
- Callaghan, G. (2008) "Evaluation and negotiated order: developing the application of complexity theory", *Evaluation*, 14(4): pp. 393-405
- Carmel, E. and Harlock, J. (2008) "Instituting the 'third sector' as a governable terrain: partnership, procurement and performance in the UK", *Policy & Politics*, Vol. 36, no. 2 pp. 155-171
- Chelimsky, E. (2007) "Factors influencing the choice of methods in federal evaluation practice", in Julnes, G. and Rog, D. (2007) (eds.) *Informing federal policies on evaluation methodology: Building the evidence base for method choice in government sponsored evaluation*, *New Directions in Evaluation* series, Vol. 113, pp. 13-33
- Cilliers, P. (2005) "Complexity, Deconstruction and Relativism", *Theory Culture and Society*, Vol. 22, No. 5, pp. 255-267
- Coburn, A. (2011) "Building Social and Cultural Capital through Learning about Equality in Youth Work", *Journal of Youth Studies* Vol. 14, No. 4, pp. 475-491
- Connell, J.P., Kubisch, A.C., Schorr, L.B., and Weiss, C.H. (eds) (1995) *New Approaches to Evaluating Community Initiatives*, vol. 1, *Concepts, Methods and Contexts*, Washington D.C.: Aspen Institute
- Connell, J. and A. Kubisch (1998) 'Applying a Theory of Change Approach to the Evaluation of Comprehensive Community Initiatives: Progress, Prospects and Problems', in K. Fulbright-Anderson, A. Kubisch and J. Connell (eds) *New Approaches to Evaluating Community Initiatives*, vol. 2, *Theory, Measurement, and Analysis*. Washington, DC: Aspen Institute
- Cooper, S. (2012) "Evaluation: ensuring accountability or improving practice", in Ord, J. (2012) (ed.) *Critical Issues in Youth Work Management*, London: Routledge
- Davies, B. (2013) "Youth work in a changing policy landscape: the view from England", *Youth and Policy*, No. 110, pp. 6-32
- Deleuze, G. (1991) *Empiricism And Subjectivity: an Essay on Hume's Theory of Human Nature*, New York: Columbia University Press
- Denzin, N.K. and M.D. Giardina (2006) *Qualitative Inquiry and the Conservative Challenge*, Walnut Creek, CA: Left Coast
- Denzin, N.K. and M.D. Giardina (2010) *Qualitative Inquiry and the Politics of Evidence*. Walnut Creek, CA: Left Coast Press
- Department for Education and Skills (2002) *Transforming Youth Work: Resourcing Excellent Youth Services*, London: HMSO
- DeStefano, L. (2000) *Evaluation As a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation*, San Francisco, CA: Jossey-Bass
- Donaldson, S.I., Christie, C.A., and Mark, M.M. (2009) (eds.) *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* London: Sage
- Feinstein, O. N. (2012) "Evaluation as a learning tool", In: Kushner, S. and Rotondo, E. (eds), *Evaluation voices from Latin America*, *New Directions in Evaluation*, No. 134, pp. 103-112
- Foucault, M. and P. Rabinow (1997) *Ethics: Subjectivity and Truth*, New York: New Press

- Gilchrist, A. (2000) "The well-connected community: networking to the edge of chaos", *Community Development Journal*, 35(3): pp. 264-275
- Gormally, S. and A. Coburn (2013) "Finding Nexus: Connecting Youth Work and Research Practices", *British Educational Research Journal* Vol. 40, No. 5, pp. 869–885
- Greene, J.C. (2009) "Evidence as 'Proof' and Evidence as 'Inkling'", in: Donaldson, S.I. et al (2009) (eds.) *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* London: Sage, pp. 153-167
- Greene, J.C., Millet, R.C., and Hopson, R.K. (2004) "Evaluation as a democratizing practice", in Braverman, M.T., Constatine, N.A., and Slater, J. K. (2004) (eds.) *Foundations and evaluation: Contexts and practices for effective philanthropy*, San Francisco: Jossey-Bass, pp. 96-118
- Gregg, M. and Seigworth, G.J., (2010) *The affect theory reader*. Duke University Press.
- HM Treasury (2011) *The Green Book: Appraisal and Evaluation in Central Government*, London: The Stationary Office
- Howe, KR (2009) "Positivist dogmas, rhetoric, and the education science question", *Educational Researcher* vol 38(6): pp. 428–440
- House of Commons Education Committee (2011) *Services for Young People: Third Report of the Session 2010 – 12 Volume 1*, HC 755-1, London: The Stationary Office Ltd.
- In Defence of Youth Work, (2012) *This is Youth Work: Stories from Practice*, London: Unison.
- Issitt, M. and Spence, J. (2005) "Practitioner knowledge and the problem of evidence based research policy and practice", *Youth and policy* No. 88, pp. 63-82.
- Julnes, G. and Rog, D. (2009) "Evaluation Methods for Producing Actionable Evidence: Contextual Influences on Adequacy and Appropriateness of Method Choice", in: Donaldson et al (2009) (eds.) *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* London: Sage, pp. 96-132
- Kuhn, T.B. (1953) *The structure of scientific revolutions*. Chicago: University of Chicago Press
- Lather, P. (2004) "This Is Your Father's Paradigm: Government Intrusion and the Case of Qualitative Research in Education", *Qualitative Inquiry*, Vol. 10, No.1, pp.15–34
- MacLure, M., (2011) "Qualitative Inquiry Where Are the Ruins?" *Qualitative Inquiry*, 17(10), pp.997-1005.
- Mark, M.M., G.T. Henry, and G. Julnes (2000) *Evaluation: An Integrated Framework for Understanding, Guiding, and Improving Policies and Programs*. San Francisco: Jossey-Bass
- Mowles, C. (2014) "Complex, but not quite complex enough: The turn to the complexity sciences in evaluation scholarship", *Evaluation*, 20(2): pp. 160-175
- Oakley, A. (2000) *Experiments In Knowing: Gender and Method in the Social Sciences*, New York: New Press
- Ord, J. (2012) (ed) *Critical Issues in Youth Work Management*, London: Routledge
- Ord, J. (2014) "Aristotles' *Phronesis* and Youth Work: Beyond Instrumentality", *Youth and Policy*, No. 112, pp. 56- 73
- Project Oracle (2014) "Standard 1: Project model and evaluation plan", *Project Oracle Children and Youth Evidence Hub*, [http://project-](http://project-children-and-youth-evidence-hub.org/)

oracle.com/uploads/files/Project Oracle Standard of Evidence level 1 - Guidance - February 2014 1.pdf (accessed: 08/08/2016)

- Scruton, R. (1999) *Spinoza*, New York: Routledge
- Sercombe, H. (2010) *Youth Work Ethics*, London: Sage
- Shaker, P. and C. Ruitenberg (2007) “‘Scientifically-Based Research’: the Art of Politics and the Distortion of Science.” *International Journal of Research & Method in Education*, Vol. 30, No. 2, pp. 207–219
- St. Pierre, E.A. (2002) “‘Science’ rejects postmodernism”, *Educational Researcher*, Vo. 31, No. 8, pp. 25-27
- Taylor, D. and Balloch, S. (eds) (2005) *The Politics of Evaluation: Participation and Policy Implementation*, Bristol: Policy Press
- Torrance, H. (2010) “Building Confidence In Qualitative Research: Engaging the Demands of Policy.” In: Denzin, N.K. and M.D. Giardina (2010) *Qualitative Inquiry and the Politics of Evidence*. Walnut Creek, CA: Left Coast Press, pp. 55-79
- Trinder, L. and S. Reynolds (2000) *Evidence-Based Practice: a Critical Appraisal*. Oxford: Blackwell Science
- Weis, L., & Fine, M. (2004). *Working method: Research and social justice*. London: Routledge.
- Weiss, C.H. (1995) “Nothing as Practical as a Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives”, in Connell, J.P., Kubisch, A.C., Schorr, L.B., and Weiss, C.H. (eds) *New Approaches to Evaluating Community Initiatives*, vol. 1, *Concepts, Methods and Contexts*, Washington D.C.: Aspen Institute
- Whyte, W. F. E. (1991). *Participatory action research*. London: Sage
- Young Foundation (2012) *An Outcomes Framework for Young People’s Services*, available at <http://www.youngfoundation.org/publications/reports/an-outcomes-framework-young-peoples-services>
- Young, K. (1999) *The Art of Youth Work*, Lyme Regis: Russell House