LSE Research Online

Flora Cornish

Evidence synthesis in international development: a critique of systematic reviews and a pragmatist alternative

Article (Accepted version)
(Refereed)

http://eprints.lse.ac.uk

Evidence synthesis in international development: A critique of systematic reviews and a pragmatist alternative

Flora Cornish

Department of Methodology, London School of Economics & Political Science, UK.

Department of Methodology

London School of Economics & Political Science

Houghton St

London WC2A 2AE

UK

Tel: +44 (0)207 955 6792

Email: f.cornish@lse.ac.uk

Flora Cornish is Associate Professor in Qualitative Research Methodology at the London School of Economics. She has researched HIV prevention for sex workers in India ethnographically. More recently she has focused on the intersection of the evidence paradigm and community mobilisation interventions.

Evidence synthesis in international development: A critique of systematic reviews and a pragmatist alternative

Systematic reviews are an instrument of Evidence-Based Policy designed to produce comprehensive, unbiased, transparent and clear assessments of interventions' effectiveness. From their origins in medical fields, systematic reviews have recently been promoted as offering important advances in a range of applied social science fields, including international development. Drawing on a case study of a systematic review of the effectiveness of community mobilisation as an intervention to tackle HIV/AIDS, this article problematizes the use of systematic reviews to summarise complex and context-specific bodies of evidence. Social development interventions, such as 'community mobilisation' often take different forms in different interventions; are made successful by their situation in particular contexts, rather than being successful or unsuccessful universally; and have a rhetorical value that leads to the over-application of positively valued terms (e.g. 'community mobilisation'), invalidating the key-word search process of a systematic review. The article suggests that the policy interest in definitive summary statements of 'the evidence' is at odds with academic assessments that evidence takes multiple, contradictory and complex forms, and with practitioner experience of the variability of practice in context. A pragmatist philosophy of evidence is explored as an alternative. Taking this approach implies expanding the definition of forms of research considered to be 'useful evidence' for evidence-based policy-making; decentralising decisions about 'what works' to allow for the use of local practical wisdom; and prioritising the establishment of good processes for the critical use of evidence, rather than producing context-insensitive summaries of 'the evidence'.

Keywords: evidence-based policy; systematic review; community mobilisation; international development; pragmatism

Introduction

Emerging in the mid-1980s, the EBP movement has gained supporters, become a commonplace in many disciplines, and become a gatekeeper for the allocation of intervention funding (Lambert 2006; Cartwright and Hardie 2012). The argument,

initially developed in the context of Evidence-Based Medicine, was that much of doctors' practice was based on tradition or preference rather than upon rigorous science (Lambert 2006). Hence, objective evidence was called for, to challenge the authority of doctors' idiosyncratic preferences, to make them more accountable, and ultimately to improve patient care by reference to the most up-to-date and authoritative evidence. A parallel argument emerged, with the move to EBP more generally, that the constraints of evidence would prevent policies from being based on ideological, moralistic, or party-political preferences, and would establish policies with a more rational basis and universal relevance. The important claims were claims about a science which is empirical (i.e. based in examination of the world rather than hearsay or ideology), and which is disinterested (i.e. independent of political influence). While such ambitions for empiricism and independence are widely understood as basic foundations of science, including social science, this paper takes issue with the model of disinterested science that EBP implies. In particular, it problematises the aspiration to abstract a 'treatment' or 'intervention' from its particular context in order to make general claims about whether the treatment should be considered to have succeeded or to have failed as an intervention worthy of repeating and funding.

Evidence Based Policy

Emerging in the mid-1980s, the EBP movement has gained supporters, become a commonplace in many disciplines, and become a gatekeeper for the allocation of intervention funding (Lambert 2006; Cartwright and Hardie 2012). The argument, initially developed in the context of Evidence-Based Medicine, was that much of doctors' practice was based on tradition or preference rather than upon rigorous science (Lambert 2006). Hence, objective evidence was called for, to challenge the authority of doctors' idiosyncratic preferences, to make them more accountable, and ultimately to

improve patient care by reference to the most up-to-date and authoritative evidence. A parallel argument emerged, with the move to EBP more generally, that the constraints of evidence would prevent policies from being based on ideological, moralistic, or party-political preferences, and would establish policies with a more rational basis and universal relevance. The important claims were claims about a science which is empirical (i.e. based in examination of the world rather than hearsay or ideology), and which is disinterested (i.e. independent of political influence). While such ambitions for empiricism and independence are widely understood as basic foundations of science, including social science, this paper takes issue with the model of disinterested science that EBP implies. In particular, it problematises the aspiration to abstract a 'treatment' or 'intervention' from its particular context in order to make general claims about whether the treatment should be considered to have succeeded or to have failed as an intervention worthy of repeating and funding.

One of the social purposes of EBP is to enable accountable decisions regarding which of a number of competing interventions should be funded (Cornish and Gillespie 2009). Should limited HIV/AIDS funds, for instance, be devoted to peer education interventions to encourage condom use, or to gender transformative self-help groups to foster women's control over their relationships, or to income generation efforts to foster women's economic security, or to male circumcision to reduce HIV transmission? To make a vast and diverse literature legible, in the EBP paradigm, technologies of 'evidence grading' and 'evidence synthesis' are employed. 'Evidence grading' systems, and 'hierarchies of evidence' offer definitions of evidence quality and associated gradings, to allow for transparent assessments of the quality of evidence, the magnitude of benefit, and the strength of recommendations for policy or practice(e.g., GRADE Working Group 2004). Randomised controlled trials (RCTs) typically sit at the top of

such systems, followed by uncontrolled studies, cohort studies, descriptive and case studies (Evans 2003; Harris et al. 2001). Experimental, quantitative evidence is prioritised, and context-specific case studies or qualitative methods are devalued. Evidence grading systems are usually applied to individual research articles. To bring those articles together, and make a statement about 'the evidence' as a whole, evidence synthesis techniques – predominantly systematic reviews – are used (Mulrow 1994). Various critiques of EBP have been made (see Lambert 2006, for an exemplary review). This paper takes issue with the summarising ambition of 'evidence synthesis' through the highly valued 'systematic review'.

Systematic reviews

A systematic review is a literature review designed to reach high levels of comprehensiveness and transparency. Waddington et al (2012, 360) characterise a systematic review as having 'a clear protocol for systematically searching defined databases over a defined time period, with transparent criteria for the inclusion or exclusion of studies, as well as the analysis and reporting of study findings.' While a systematic review could be employed to answer a variety of questions, the most common is the effectiveness question: What is the evidence that a particular intervention produces particular desired outcomes? (White and Waddington 2012). Originally established in medical research, their perceived value as tools of EBP has led to increasing enthusiasm for systematic reviews in a range of applied social science areas, including international development, and management.

Proponents of systematic reviews give two main arguments for their use: eliminating bias and providing clarity. Compared with cherry-picking studies to fit a pre-conceived policy preference, the transparent protocol of a systematic review is intended to minimise bias (DFID Research and Evidence Division 2010; Petticrew and

Roberts 2006). Systematic reviews provide a synthesis of complex bodies of evidence, overcoming 'information overload' (Petticrew and Roberts 2006; Waddington et al. 2012). Introducing its own systematic review program, the DFID Research and Evidence Division (2010, 2) writes that systematic reviews present 'a clearer and more consistent picture of the body of evidence, [...one] that is accessible and easily understandable to non-experts'.

Few social scientists would disagree that informed, appropriate use of social scientific research is good for policy-making, and indeed that listening to systematic reviews rather than relying on individual primary studies, ideology, or hearsay is a good idea. This paper does not argue against the idea of systematically collating evidence, critically appraising it, and bringing it to bear on policy decisions. It does argue, from a perspective rooted in ethnographic engagement in the context-specificities of practice, and a reflexive awareness of the rhetorical significance of claims about interventions and practice, that the impetus to produce an authoritative, clear statement of 'the evidence' leads to the misuse of systematic reviews to provide false certainty. This argument is grounded in practical experience of conducting a systematic review (Cornish et al. 2014).

Case study: Systematic review of evidence on community mobilisation and HIV intervention

With the expectation (in hindsight, a naïve one), that a systematic review of the evidence for community mobilisation would provide useful information, or at least useful rhetoric, a group of colleagues and I recently undertook a systematic review of the evidence for the impact of community mobilisation within the HIV response in low and middle income countries (Cornish et al. 2014). 'Community mobilisation' is a term that refers to efforts to build and capitalise on community strengths, networks, resources

and agency. In the HIV/AIDS field, community mobilisation activities are funded as part of interventions, in the interest of achieving improvements in terms of HIV-related outcome measures. The 'communities' that are mobilised often have interests in broader social change agendas such as tackling discrimination or securing their rights as citizens. 'Community mobilisation' is conceptualised in different ways, sometimes as 'mobilising' communities in something akin to a social movement, or collective action to tackle HIV/AIDS, sometimes in a more instrumental way as 'mobilising' community resources in the service of HIV/AIDS interventions. The term grew out of the discourse of 'community participation' popular in the 1990s. While it was originally used to signify an approach to community work that had a politicised angle, initiating collective action through 'mobilisation', it has increasingly been used to refer to any activity that involves bringing community members together (Cornish et al., in prep). Our systematic review sought to review 'authentic' community mobilisation, not 'tokenistic' community mobilisation, but this task proved harder than it appeared at first sight. The review considered impacts on measures of public health interest, including biomedical outcomes (e.g. STI prevalence), behavioural outcomes (e.g. reported condom use), and micro-social outcomes (e.g. collective efficacy).

Our searching and screening process identified 20 relevant studies for inclusion in the review. Among these studies, some reported interventions conducted with 'risk groups' (9 interventions with sex workers; 2 with men who have sex with men, 1 with heterosexual men who drink alcohol); some with a more 'general' population (4 with young people; 4 with adults in a geographically-bound community). Our initial conceptualisation of the scope of the review was to synthesise the evidence for 'community mobilisation' in general. However, one of the findings to emerge was that the evidence was more positive for the 'at risk groups' than for the 'general population'.

This difference problematized the impetus to review 'the evidence in general', if the evidence was different for different groups.

The findings of the review, in common with so many reviews of social interventions (Pawson and Tilley 1997; MacLure 2005) were 'inconclusive'. Even to summarise the findings proves difficult. Guidelines for systematic reviewers caution, rightly, against 'vote-counting' (that is, counting the numbers of studies producing positive, inconclusive and negative results) as a way of summarising the evidence, given that the different sample sizes, research designs, and measures used mean that the findings of studies cannot simply be added together to give a meaningful answer (Petticrew and Roberts 2006). Nonetheless, vote-counting can illustrate the variability in findings that emerged. In the studies with sex workers, 3 studies found improvements in biomedical outcomes, and 3 found none. Reported condom use tended to increase, but within studies, findings were patchy, for instance improving with paying partners but not with non-paying partners. In studies with youth and 'general population' adults, the findings were less positive, but still patchy. On biomedical measures, one study reported improvements, 5 found no effects, and 2 found increases in STI/HIV. Overall, we described the findings of the review as indicating that, among at risk groups (particularly sex workers), the evidence was 'somewhat consistent, indicating a tendency for positive impact' (Cornish et al. 2014, 1), and that among youth and general communities 'the evidence remains inconclusive'.

Substantively, then, it proved practically impossible to come to a reasonable and simple conclusion about whether 'community mobilisation' works or not. We concluded that this was not because 'community mobilisation' has 'failed' the EBP test. Rather, our conclusion was that understanding of 'evidence', as conceptualised under the regime of EBP, has failed to do justice to community mobilisation. In reflecting on

the reason for our 'inconclusive' findings, we identified three main problems with 'the evidence' as it stands. The first concerned operationalization of 'community mobilisation'. Despite the term having been introduced to distinguish it from tokenistic forms of 'community participation', 'community mobilisation' itself had become instrumentalised and tokenistic (Trickett 2011). In many interventions, 'community mobilisation' was described as a process of 'consulting' with 'the community' or 'community leaders', or 'sensitising' the community (e.g. Cowan et al. 2008; Jewkes et al. 2006). In such cases, 'mobilisation' often meant simply recruiting people to turn up. Thus, what is labelled 'community mobilisation' in this literature is not necessarily good material with which to test the usefulness of mobilising communities in a process of collective action.

The second reason for inconclusive evidence was again a question of the design of community mobilisation interventions. Many of the interventions failed to engage with the broader social and political context and power relations that structure health in very disadvantaged communities. Contemporary understandings of community mobilisation emphasise that communities alone rarely have the power to make the social changes needed to sustain healthy behaviour, and hence, that community mobilisation needs to be accompanied by efforts to engage powerful stakeholders in the pursuit of structural change (Campbell and Cornish 2010; Cornish and Campbell. 2010). In programmes involving youth and general communities, there was evidence of limited efforts to engage the broader community. Where efforts were made to engage groups beyond the target group, this often had the limited aim of enabling the diffusion of health-related knowledge rather than engaging them in transformative change.

Thus, the systematic review process de-contextualises the reports of intervention (i.e. the use of the label 'community mobilisation' in a published article) from the

practical activities of the intervention (e.g., the agency accorded to communities versus intervention leaders; the relationships among community members; the freedoms of the community members to adapt the project goals; the scope of the intervention to address structural as well as micro-social levels). The materials with which a systematic review operates are published academic articles, that is, texts, not the actualities of practice. The relation of those texts to concrete practices is not clear, and the rigorous and replicable application of keyword searches does not allow space for critique of how community mobilisation was implemented in any particular case.

Thirdly, reflecting on the very uneven nature of the findings, we suggested that the goal of providing an over-arching statement of 'the evidence' for community mobilisation may itself be misguided (Seckinelgin 2010). We observed a different pattern of findings for sex workers and for youth and general communities. In some studies, there appeared to be impacts on condom use with some types of partners, but not others. The IMAGE study found impressive effects on intimate partner violence (widely argued to be a likely contributor to HIV transmission), but no effects on HIV incidence. Based on earlier positive results from the Sonagachi Project in Kolkata (Jana et al. 2004; Jana and Singh 1995), replications were implemented in Brazil (Lippman et al. 2012; Kerrigan et al. 2008) and elsewhere in India (Swendeman et al. 2009; Basu et al. 2004) but to less or no positive effect. Furthermore, evaluations of the Avahan programme disaggregated aspects of community mobilisation and their impact on a host of measures of condom use in a variety of settings, finding some significant relationships at a fine-grained level, but little consistency (Saggurti et al. 2013; Blanchard et al. 2013; Guha et al. 2012).

Systematic review: Three problems for context-sensitive processes

The experiences and reflections reported above from our own case study are not unique.

As EBP has gained in stature, so have critical reflections. Three particular problems for context-sensitive processes are clear.

Restricted definition of 'evidence'

In a paper compellingly titled 'What is the evidence that there is no evidence?', Hakan Seckinelgin (2010) provides a detailed critique of the conclusion of 'no evidence' reached by a systematic review of evidence on the relation between conflict and HIV/AIDS. Seckinelgin questions the types of data which are deemed to comprise evidence, namely quantitatively sophisticated studies designed to eliminate 'bias'. Unsurprisingly, there are limited numbers of such studies, and implementing a 'strict' definition of evidence discounts the vast majority of informative studies and theories that can contribute to building sound conclusions, but are not admissible, under the EBP regime, as 'evidence'. The general argument is that the definition of 'evidence' here fails to meet the needs of evidence users, compelling them to make less informed decisions than they might otherwise do.

Ignoring context and complexity

Pawson and Tilley's (1997) book 'Realistic Evaluation' remains a classic on the topic of context. They note that the conclusion of most efforts to summarise or review the evidence on social interventions is that the research is 'inconclusive', with often a mixed bag of results, some positive, some negative, some indicating no effect. Pawson and Tilley's main argument is that, rather than seeking blunt statements about whether an intervention works or not, researchers ought to be conceptualising the conditions under which a particular mechanism is more or less likely to work. An intervention might work in the presence of certain supportive conditions, but not without them. It might work for people with certain characteristics but not others. Kippax (2012) argues

that herein lies the flaw in applying the methods privileged by EBP to the evaluation of social interventions. She points out the distinction between efficacy (does a treatment work under ideal conditions, all things being equal?), and effectiveness (does a treatment work in the real world?). HIV medicines work well under ideal conditions, but in the absence of sufficient food, they do not (Seckinelgin 2012). As Mallett et al. (2012, 452) argue: '"cutting out the noise" risks missing the point in international development research (and the social sciences more broadly), where context is the primary consideration'.

Academic authors writing about evidence and systematic reviews typically employ important caveats, including: the importance of exercising caution in generalising across contexts, the need to draw on local expertise to interpret evidence in relation to a particular policy context, the inconclusiveness of much of the research literature, the value of a range of forms of evidence, or that evidence should be seen as a contributor (among other sources) to EBP, rather than being considered the sole pillar of EBP (e.g. White and Waddington 2012; Shepperd et al. 2009). However, one of the main rationales given for systematic reviews in the context of policy-making is precisely to deal with 'information overload' (Petticrew and Roberts 2006, 22) or the 'bewildering amount of primary studies' (DfID Research and Evidence Division 2010, 2).

MacLure (2005), based on her review of 30 systematic reviews, comments on the pressures on systematic review authors to offer recommendations, even if the findings are inconclusive. She observed that authors often acknowledged the lack of a secure evidentiary base, stated that it is not possible to draw implications confidently, and yet, in every review, they went on to put forward findings, recommendations and/or implications. In reporting our own review, we felt it would make an unsatisfying read to

conclude that the findings were inconclusive. Indeed a peer-reviewer explicitly recommended that we include useful 'lessons learned' for practice. Accordingly, we concluded that community mobilisation is more likely to succeed with meaningful identity groups than with abstract 'population' groups. This interpretation can be supported by the review, but of course it also depends on a theory of community (as mutual identification), and goes beyond the data presented in the review. While an academic audience might tolerate a conclusion that no firm conclusion can be drawn, to be a useful policy tool, a systematic review needs to offer endorsement of one course of action or another. This is one source of the simplification and erasure of nuance that happens when variable scientific knowledge is subjected to systematic review.

Aspiration to objectivity / Denial of reflexivity

The EBP movement was premised on the value of striving for objectivity (rather than the subjective preferences of doctors, or the political leanings of politicians). Eliminating 'bias' on the part of the reviewer (along with replicability and rigour) is one of the three main planks in the argument for systematic reviews (e.g. DFID Research and Evidence Division 2010; Petticrew and Roberts 2006; White and Waddington 2012). Counter to this ambition, it has been argued that critical specialist expertise is required for a meaningful systematic review (MacLure 2005). When the 'treatment' being evaluated is fuzzy or complex, the need for judgment becomes all the more acute. In conducting our systematic review, we knew we wanted to define community mobilisation to exclude the most instrumental operationalisations, as far as possible. This impetus came from our own experiences as researchers in the field, having observed practices labelled as 'community mobilisation' which did not fit with our own understandings of that term. Hence, we were suspicious of articles reporting community mobilisation being used simply, for example, to recruit men for circumcision. Articles

reporting findings of quantitative studies often presented only very sketchy accounts of their community mobilisation activities. Where possible, we combed qualitative accounts of the trials. Inevitably, we had to use our judgement to decide whether 'community meetings', 'community theatre', or 'consultations' should be classified as 'community mobilisation' or not. Indeed, we had long discussions about what we meant by 'real' community mobilisation, and how to write in to our inclusion and exclusion criteria the human agency and empowerment that we considered to distinguish 'community mobilisation' from instrumental peer education.

Systematic review guidelines implicitly acknowledge the need for expert judgement, recommending that in the process of identifying potential articles for review, as well as searching databases, it is appropriate to contact researchers in the field to aid identification of relevant studies (Petticrew and Roberts 2006; Gough et al. 2012). We followed this advice, and included the important IMAGE study (Pronyk et al. 2006), after consultation with colleagues, even though it did not use our designated key-words in its title or abstract. One of the peer-reviewers of our submitted article suggested a further study that ought to be included. Widening our search in this way enabled us to cover a greater literature, but also indicates the contingency involved in the inclusion of articles. Moreover, we drew on our ethnographic experiences and networks, as scientists and practitioners, in the worlds of both intervention implementation and social science. We knew that the Sonagachi project was famed as a 'successful' community mobilisation project, but that it was also controversial. And, in particular, we were attuned to the over-use of the term 'community mobilisation', having encountered community mobilisation in action, rhetorical claims about the use of community mobilisation, and dissatisfaction among NGO workers we had spoken to

in India, who complained about their NGOs misrepresenting their instrumental top-down work as the more politically desirable 'community mobilisation'.

Here we encounter again a familiar complexity of the social sciences over the natural sciences, namely the reflexivity of human categories, or in other words, the fact that scientific terms are part and parcel of everyday life and everyday politics, and cannot be treated as insulated from those realms. 'Community mobilisation' is not a term with an objective referent. Moreover, it is a term that has rhetorical value (e.g. claiming that this intervention was ethically sound, or adhered to contemporary best practice). Given this, it is foreseeable that the term will be over-applied. Projects survive by claiming to implement best practice and by claiming success (Mosse 2005; Cornish and Ghosh 2007). Decontextualising the label 'community mobilisation' (by taking authors' claims to community mobilisation at face value) fails to acknowledge that 'community mobilisation' is a construction; the deployment of the term is often strategic, a rhetorical move rather than a representation of reality (see also Greenhalgh and Russell 2006). This suggests that it is not at all clear what exactly is being summarised and reviewed, when one reviews literature claiming to implement 'community mobilisation'.

Together, these three problems deeply undermine the capacity of general statements about 'the evidence' for 'community mobilisation' to be meaningful. A statement such as 'there is a tendency for positive impact of community mobilisation' can be grammatically correct and methodologically justified, based on a rigorous and replicable systematic review method. But the term 'community mobilisation' does not refer to a single and stable entity. It is implemented differently in different interventions, is used to achieve different goals depending on situational demands, and may work through different mechanisms in different contexts. For statements about

context-sensitive processes such as community mobilisation to be meaningful, they have to be unpacked back into particular concrete settings and practices. Some people are mobilising, using certain participatory practices, to achieve particular goals. While such an argument for the specific and concrete seems at odds with the implicit universalising philosophy of science behind systematic reviews, it is not an anti-scientific argument, as the following section elaborates.

An alternative philosophy of evidence: Pragmatism

This section turns to elaborating a perspective on evidence based in a pragmatist philosophy of science. What is notable about this philosophy is that, despite pragmatism's test of truth, just like the EBP movement, being motivated by the question of 'what works', its conceptualisation of science is almost diametrically opposed to that implied in EBP. Specifically, the aspiration of EBP to objectivist, generalisable truths is opposed by a pragmatist version of science as an art of experience-based, context-specific, gradual improvement. To elaborate this perspective, the section explores four important principles of a pragmatist theorisation of science: the primacy of practice; the tool metaphor of knowledge; the focus on particulars over universals, and the experimentalist ethos of bringing about social change.

Primacy of practice

The defining characteristic of pragmatism is its privileging of practice as the bedrock and testing-ground of knowledge (Brinkmann 2013). Pragmatists argue that we have no grounds, and little to gain by trusting either in the realist aspiration for our concepts to match an unobservable reality, or the idealist advancement of the power of reflective thought to reach a purer truth than the messy truths of practical activity (Rorty 1999; Cornish and Gillespie 2009). Practice is empirical, experiential, contextual and

historical. Being empirical, practice is what people do in a messy world. Here, the empirical has priority over abstract, universalised theorising. Pragmatists argue that we have no cause to doubt the reality of practice, in the way that we can doubt the reality of abstracted claims. Being experiential, practice links human subjectivity to worldly objects. Knowledge has to be in the meeting of human interests and worldly materials; it is not objective or cleansed of human interests or values. Being contextual and historical, practice is a response to the multiple demands of a particular place (it cannot be universal), and evolves over time, responding to what has gone before, and bringing about what comes next (it is not knowledge separate from the world, but it is world-constituting practical knowledge).

The primacy of practice has had a resurgence in the recent 'practice turn' in social sciences (Schatzki, Knorr Cetina and von Savigny 2001). Practice theorists privilege the local, particular and messy, over the presumption to sketch grand universalising theories (e.g. de Certeau 1984). In the context of theorising community health action, Nolas (2014), drawing on Deleuze and Guattari (1984), describes the practice turn as prioritising 'journeys' over 'maps'. Maps represent idealised, abstracted intentions. In the context of interventions, maps typically represent a deterministic theory, of how to bring about personal or community change. Journeys, by contrast, are thick, contentful, peopled stories, reporting the messy, troubling, relational and contingent hard graft of working to produce personal or community change. Practice theorists do not deny the contingencies, unpredictabilities, and context-specificities of practice. Rather, they assume that the empirical world has this 'messy' character, and thus that the work of social science to understand, explain, or change must have its foundations firmly rooted in skilled working with 'mess' rather than in aspirations to decontextualized universalising principles. As the following sections elaborate,

prioritising practice over grand theory has implications for the understanding of knowledge, evidence, science, and intentional change.

Truths made, not found

The key pragmatist intervention in epistemological debates has been to change the question that we ask about knowledge, from the commonsense and realist question of 'Does this knowledge accurately reflect the underlying reality?' to the pragmatist question of 'Does this knowledge serve our purposes?' (Rorty 1999). Introducing the criterion of 'serving our purposes' puts interests (purposes) at the heart of determinations of truth. It also prioritises 'what works' as a criterion of 'truth' (rather than reference to an ultimate reality or an encompassing, coherent theory), echoing the EBP catchphrase. But it is even more radical. On questions of validity, Rorty counsels us to: 'stop worrying about whether what one believes is well grounded and start worrying about whether one has been imaginative enough to think up interesting alternatives to one's present beliefs' (Rorty 1999). Rorty argues that humans are engaged in producing their own futures, and that we should give greater attention to imagining and pursuing the future we would prefer, rather than trying to 'discover' truths about the present. He considers literature as the pinnacle of human crafts, for its capacity to envision alternative futures. Such an argument would problematize the hierarchies of knowledge espoused by the EBP movement, and value theorising, critique, prefigurative politics, illuminative case studies, and any other methods that provoke novel and liberating figurations.

Understanding truths as made in practice has a further critical consequence for discussions of 'evidence'. Human practice is intelligent, reflective and productive. It does not simply reflect reality but makes reality. From this perspective, community mobilisation is not an inert object whose effectiveness can be measured. Successes and

failures are made in practice, as a result of the intersection of people, activities, resources and constraints. To some extent, community mobilisation can be made to work. Its success partly depends on the degree to which we value community mobilisation enough to invest enough resources in it and commitment to it.

Particulars, universals and judgements in context

From a pragmatist, and practice-theoretical point of view, the tools for action are not theories with universal claims, but expertise grounded in experience (Flyvbjerg 2001). In his book 'Making Social Science Matter', Flyvbjerg seeks to revive the status of 'phronesis' or practical wisdom, as a valued accomplishment of social science. Because human action is so context-dependent, and thereby not deterministic, Flyvbjerg argues that a good social scientist becomes an expert in their field – just as surgeons or football players become experts – through practical engagement in multiple, varying instances of their phenomenon in its natural context. Whereas the claims of social science to generality have been argued on the basis of universalising abstract theory, or on the basis of generalising from a sample to a population, Flyvbjerg's argument is that it is through expert judgement in context that experience from one setting is made relevant and applicable to a new setting. Relatedly, in a study assessing evidence in an educational setting, Greenhalgh et al. (2003) argue that qualitative studies of student and staff experience played a crucial role in determining the validity and transferability of existing evidence.

'Mess' and contingency do not imply unknowable, unpredictable, or paralysing chaos. But they do imply complexity, and interdependence of phenomena and their context. Dewey (1925) argues that science is an art, where art is understood as speaking both to the contingencies and to the regularities of reality. The art of social science consists in the scientist's expertise in diagnosing, analysing, understanding and

improving a particular local problem, in the light of past experience. According to Dewey (1925), what is needed is 'discriminating judgement by methods whose consequences improve the art' (p.383). The word 'judgement' signifies a skilled human observer, who weighs the evidence, and comes to a judgement about the case, and what is likely to work. A judgement is not a statement of fact, but a considered outcome of weighing indeterminate evidence, building on past experience. A judgement is made about a specific case, and reflects its context.

In the context of EBP, valuing expert judgement in this way returns to us the risk, which EBP claimed to counter at the outset, of an over-reliance on the peculiarities of individual clinicians. One response to this risk might be to maintain the use of judgement but to make it more accountable, by being more social, as I shall suggest, below.

Change through gradual experimentalism

The final principle I wish to address is that of experimentalism as the process of social and political change. Dewey concerned himself with the question of how transformational social change could come about without the violence and disruption of revolution. He sought out ways that democracy could bring about transformed societies, but incrementally and organically. Roberto Unger (2007) has taken up this project more recently, arguing that contemporary societies should concern themselves with creating organisational systems that allow for continual reinvention. He writes:

> our societies and cultures may be so arranged as to facilitate and to organize
> their own piecemeal, experimental revision. We then shorten the distance
> between routine moves within a framework and exceptional moves about the
> framework; we experience the latter as a direct and frequent extension of the

former. As a result, we denaturalise society and culture: we unfreeze them (Unger 2007, 7).

This is another argument in favour of local, incremental adaptation, in response to context, as opposed to the ambition of EBP to identify an intervention that 'works' which can then be 'scaled up' and 'rolled out'.

The vision of how we organise our societies to try to improve them, espoused here, is to enable local, organic and piecemeal change. Two important implications for the policy-making process follow. The first is an argument that we should focus on creating good processes for collective decision-making and collection action, rather than focusing on achieving a pre-determined outcome (Brinkmann 2013). The second is a consideration of interventions and their outcomes as part of a cycle of ongoing invention. In this cycle, much trial-and-error learning takes place. 'Failure' is part of the learning process, and informs improvements. 'Failed' interventions may not be failures at all, but part of a learning cycle, creating knowledge to be put to use in the next experiment in the cycle.

Towards pragmatist versions of 'evidence'

What might 'evidence' look like in a policy-making process designed by a pragmatist? Pragmatism does not indicate that 'anything goes', or that contingency and localism imply that there is nothing more than trial and error learning to be done. I make three preliminary suggestions for directions which a pragmatist version of 'evidence' might imply.

Firstly, the approach which I have outlined calls for a broadening of the understanding of 'evidence' beyond the prioritisation of systematic review and RCT. Local case studies of intervention processes in context, theorisations of practice,

experimentation with novel intervention processes, perspectives of local people – these are all sources of information which do not contribute to EBP as currently defined, but which build valued intellectual resources for informing action. In most areas of social action, our knowledge does not allow accurate prediction. Rather than excluding potentially informative studies because they do not fit the highly restrictive methodological criteria traditionally applied in systematic reviews (MacLure 2005), a good pragmatist use of evidence would make use of the variety of research available. Given the exponential expansion of the research literature over the last several decades, can it really be true that a systematic review in any area can either meaningfully be based on less than ten studies as is common (MacLure 2005), or that 'more research is needed' (Billig 2013; Greenhalgh 2012 'Less research is needed'

http://blogs.plos.org/speakingofmedicine/2012/06/25/less-research-is-needed/)?

Secondly, evidence synthesis is possible, but different, from a pragmatist point of view. The process of weighing this evidence, if we accept the arguments of Dewey and Flyvbjerg in favour of expert human judgement, would place greater reliance on skilled consideration of evidence, assessed with reference to experience in context. The 'diagnostic approach' to intervention design is one example of such a process (Gupta et al. 2008; Mannell, Cornish & Russell, 2014; Parkhurst 2013). Following this approach, rather than assuming that evidence is easily transferred across contexts, the task of intervention design is to critically and sensitively assess which causal mechanisms are most important in a given context. An analogy to the use of evidence in a court, in which experienced judges weigh a variety of sources of information, might also be a helpful model. Making this process social, and particularly, involving diverse voices in this new court (including representatives of local communities; clinicians; scientists;

economists; civil servants; etc) is one way to counter the risks attached to claims of expert judgement: that they may reflect personal whim or political agenda.

Thirdly, regarding the kinds of initiatives to be supported by policy, the pragmatist approach suggests focusing on processes rather than pre-determined outcomes. Policy might support local experimentation by reducing the emphasis on reaching pre-determined, national-level objectives. It might initiate processes for the improvement of practice, rather than prescribing specific practices. In the field of 'improvement science' for instance, Aveling et al. (2012) report on 'clinical communities' as a means of enabling clinicians to work collectively to improve their practices. The focus is on the means of working together to change practices, rather than on specific outcomes.

Conclusion: Evidence in context

This article has problematised the idea that it is possible to make a summary statement of whether an intervention 'works' or not, by means of a systematic review. It has argued against making universalising statements about 'what works' by arguing for the importance of local context and local judgement to answer the more relevant question of 'what works for this particular time, place, and these goals'. Overall, the article has argued that evidence is not abstract, disinterested or disembodied. To be 'evidence', knowledge has to be used by an evidence user. Understandings of 'evidence' are themselves located, embodied in particular persons with particular histories, interests and social locations. What is deemed 'evidence' in different contexts is consonant with the human practices of different social roles. Administrators are required to made decisions about the allocation of resources. They are increasingly called to account for their decisions in a transparent manner. Reference to systematic reviews of 'the evidence' supports the goals of displaying fairness and transparency (Cornish and

Gillespie 2009). In contrast, NGO workers and communities may seek to understand local practices and create interventions that are plausible, feasible, and empowering locally. In this context, reference to decontextualized 'evidence' provides little more than a suggestion as to what has happened elsewhere, which might be made relevant to local needs, and which might (or might not) feasibly mobilise sufficient interest and commitment to make it worth implementing. Research scholars might contribute to either of these sets of interests, or to other interests in advancing understandings of the policy-practice knowledge system as a whole, or to advance theory or critique. Social science need not prioritise abstract universalising statements. While such statements may be the concern international donors and policy-makers, scientific research can be more nuanced (Behague et al. 2009). A pragmatist science of evidence would begin with the 'mess' of practice and seek to understand complexity and contingency rather than efface it. For those who work with people and communities, in the complexity of practice, the aspiration to produce simple summaries of evidence appears puzzling. From a perspective rooted in local realities, aspirations to draw on a diversity of forms of knowledge; to draw on practical wisdom to design interventions suited to contexts; and to develop models for making good context-sensitive judgements appear more fitting priorities.

References

Aveling, E.-A. et al. (2012). "Quality Improvement through Clinical Communities: Eight Lessons for Practice." Journal of Health Organization and Management, 26(2): 158-174.

Basu, I., Jana, S., Rotheram-Borus, M.J., et al. 2004. "HIV Prevention among Sex Workers in India." Journal of Acquired Immune Deficiency Syndrome 36(3): 845-52.

Billig, M. 2013. Learn to Write Badly: How to Succeed in the Social Sciences. Cambridge: Cambridge University Press.

Blanchard, A.K., Mohan, H.L., Shahmanesh, M., et al. 2013. "Community Mobilization, Empowerment and HIV Prevention among Female Sex Workers in South India." BMC Public Health 13(1): 234.

Brinkmann, S. 2013. John Dewey: Science for a Changing World. London: Transaction Publishers.

Campbell, C. and Cornish, F. 2010. "Towards a 'Fourth Generation' of Approaches to HIV/AIDS Management: Creating Contexts for Effective Community Mobilization." AIDS Care, 22, 1569-1579.

Cartwright, N. and Hardie, J. 2012. Evidence Based Policy: A Practical Guide to Doing it Better. Oxford: Oxford University Press.

Cornish, F., Dashtipour, P., Priego-Hernandez, J., and Mannell, J. (in prep). What is community mobilization? Five different approaches from an international study of community mobilization in the HIV response

Cornish, F. and Campbell, C. (Eds.) 2010. "The Social Context of Community Mobilization: Foundations for Success or Failure." AIDS Care, 22 (Suppl. 2).

Cornish, F. and Ghosh, R. 2007. "The Necessary Contradictions of 'Community-Led' Health Promotion: A Case Study of HIV Prevention in an Indian Red Light District." Social Science & Medicine 64, 496-507.

Cornish, F. & Gillespie, A. 2009. "A Pragmatist Approach to the Problem of Knowledge in Health Psychology." Journal of Health Psychology, 14, 1-10.

Cornish, F., Priego Hernandez, J., Campbell, C., Mburu, G. and McLean, S. (2014) "Impact of Community Mobilisation on HIV Prevention in Middle and Low Income Countries: A Systematic Review and Critique." AIDS & Behavior DOI 10.1007/s10461-014-0748-5

Cowan, F.M., Pascoe, S.J., Langhaug, L.F., et al. 2008. "The Regai Dzive Shiri Project: A Cluster Randomised Controlled Trial to Determine the Effectiveness of a Multi-Component Community-Based HIV Prevention Intervention for Rural Youth in Zimbabwe–Study Design and Baseline Results." Tropical Medicine & International Health 13(10): 1235-44.

De Certeau, M. 1984. The Practice of Everyday Life. Berkeley: California University Press.

Deleuze, G. and Guattari, F. 1987. A Thousand Plateaus. London: Continuum.

Dewey, J. 1925. Experience and Nature. New York: Dover.

Evans, D. 2003. "Hierarchy of Evidence: A Framework for Ranking Evidence Evaluating Healthcare Interventions." Journal of Clinical Nursing, 12, 77-84.

Flyvbjerg, B. 2001. Making Social Science Matter: Why Social Inquiry Fails and How it Can Succeed Again. Cambridge: Cambridge University Press.

GRADE Working Group. (2004). Grading Quality of Evidence and Strength of Recommendations. British Medical Journal, 328, 1490-1494.

Guha, M., Baschieri, A., Bharat, S., et al. 2012. "Risk Reduction and Perceived Collective Efficacy and Community Support among Female Sex Workers in Tamil Nadu And Maharashtra, India: The Importance of Context." Journal of Epidemiology & Community Health 66(Suppl. 2): ii55-ii61.

Gupta G.R., Parkhurst J.O., Ogden J.A., Aggleton P., and Mahal A. 2008. "Structural Approaches to HIV Prevention." Lancet, 372(9640):764–75.

Greenhalgh, T., and Russell, J. 2006. "Reframing Evidence Synthesis As Rhetorical Action in the Policy Making Drama." Healthcare Policy, 1(2): 34-42.

Greenhalgh, T., Toon, P., Russell, J., Wong, G., Plumb, L., and Macfarlane, F. 2003. "Transferability of Principles of Evidence Based Medicine to Improve Educational Quality: Systematic Review and Case Study of An Online Course in Primary Health Care." British Medical Journal, 326(7381): 142-145.

Harris, R. P., Helfand, M., Woolf, S. H., Lohr, K. N., Mulrow, C. D., Teutsch, S. M., et al. 2001. "Current Methods of the US Preventive Services Task Force: A Review of the Process." American Journal of Preventive Medicine, 20(3S1), 21-35.

Jana, S., Basu, I., Rotheram-Borus, M.J., et al. 2004. "The Sonagachi Project: A Sustainable Community Intervention Program." AIDS Education and Prevention, 16(5): 405-414.

Jana S. and Singh S. 1995. "Beyond Medical Model of STD Intervention—Lessons from Sonagachi." Indian Journal of Public Health, 39: 125-131.

Jewkes, R., Nduna, M., Levin, J., et al. 2008. "Impact of Stepping Stones on Incidence of HIV and HSV-2 and Sexual Behaviour in Rural South Africa: Cluster Randomised Controlled Trial." British Medical Journal 337: a506

Kerrigan, D., Telles, P., Torres, H., et al. 2008. "Community Development and HIV/SIT-Related Vulnerability among Female Sex Workers in Rio de Janeiro, Brazil." Health Education Research 23(1): 137-145.

Lambert, H. 2006. Accounting for EBM: Contested Notions of Evidence in Medicine. Social Science & Medicine, 62(11): 2633- 2645.

Lippman, S.A., Chinaglia, M., Donini, A.A., et al. 2012. "Findings from Encontros: A Multilevel STI/HIV Intervention to Increase Condom Use, Reduce STI And Change the Social Environment among Sex Workers in Brazil." Sexually Transmitted Diseases 39(3): 209-16.

MacLure, M. 2005. "Clarity Bordering on Stupidity: Where's the Quality in Systematic Review?" Journal of Education Policy, 20: 393–416.

Mallett, R., Hagen-Zankerb, J., Slater, R. and Duvendack, M. 2012. "The Benefits and Challenges of Using Systematic Reviews in International Development Research." Journal of Development Effectiveness, 4(3): 445-455.

Mannell, J., Cornish, F. and Russell, J. (2014). Evaluating social outcomes of HIV/AIDS interventions: a critical assessment of contemporary indicator frameworks. Journal of the International AIDS Society, 17: 19073.

Mosse, D. 2005. Cultivating Development: An Ethnography of Aid Policy and Practice. London: Pluto Press.

Mulrow, C.D. 1994. "Rationale for Systematic Reviews." British Medical Journal, 309, 597-598.

Nolas, M. 2014. "Towards a New Theory of Practice for Community Health Psychology." Journal of Health Psychology, 19(1): 126-136.

Parkhurst, J. O. 2013. Structural Drivers, Interventions, and Approaches for Prevention of Sexually Transmitted HIV in General Populations: Definitions and an Operational Approach. Arlington, VA: USAID's AIDS Support and Technical Assistance Resources, AIDSTAR-One, Task Order 1, and London: UKaid's STRIVE research consortium.

Pawson, R. and Tilley, N. 1997. Realistic Evaluation. London: Sage.

Petticrew, M. and Roberts, H. 2006. Systematic Reviews in the Social Sciences: A Practical Guide. Oxford: Blackwell Publishing.

Pronyk, P.M., Hargreaves, J.R., Kim, J.C., et al. 2006. "Effect of a Structural Intervention for the Prevention of Intimate-Partner Violence and HIV in Rural South Africa: A Cluster Randomised Trial." Lancet 368: 1973-1983.

Rorty, R. 1999. Philosophy and social hope. London: Penguin.

Saggurti, N., Mishra, R.M., Proddutoor, L., et al. 2013. "Community collectivization and its Association with Consistent Condom Use and STI Treatment-Seeking Behaviors among Female Sex Workers and High-Risk Men who have Sex with Men/Transgenders in Andhra Pradesh, India." AIDS Care 25(suppl1): S55-S66.

Seckinelgin, H. 2010. "What is the Evidence that there is No Evidence? The Link Between Conflict And HIV/AIDS." European Journal of Development Research, 22(3): 363-381

Seckinelgin, H. 2012. "The Global Governance of Success in HIV/AIDS Policy: Emergency Action, Everyday Lives and Sen's Capabilities." Health and Place, 18(3): 453-460.

Shepperd, S., Lewin, S., Straus, S., et al. 2009. "Can We Systematically Review Studies That Evaluate Complex Interventions?" PLoS Medicine 6(8): e1000086.

Swendeman D, Basu I, Das S, et al. 2009. Empowering Sex Workers in India to Reduce Vulnerability to HIV and Sexually Transmitted Diseases. Social Science & Medicine 69: 1157-66.

Trickett E.J. 2011. "Community-Based Participatory Research as Worldview or Instrumental Strategy: Is it Lost in Translation(Al) Research?" American Journal of Public Health, 101(8): 1353-55.

Unger, R. 2007. The Self Awakened: Pragmatism Unbound. Cambridge, MA: Harvard University Press.

Waddington, H., White, H., Snilstveit, B. Hombrados, J. G., Vojtkova, M., Davies, P. Bhavsar, A. et al. 2012. "How To Do a Good Systematic Review of Effects in International Development: A Tool Kit." Journal of Development Effectiveness, 4(3):359-387.

White, H. and Waddington, H. 2012. "Why Do We Care About Evidence Synthesis? An Introduction to the Special Issue on Systematic Reviews." Journal of Development Effectiveness 4(3): 351-358.