

Where's the 'Evidence'? Reflecting on Monitoring and Evaluation within Sport-for-Development.

Ruth Jeanes (Monash University)

Iain Lindsey (Edge Hill University)

Abstract

Purpose

The purpose of this chapter is to critically examine the extensive calls for enhanced evidence within the sport-for-development field. The chapter questions whether these are appropriate and realistic.

Design/Methodology/Approach

The chapter utilises current literature to deconstruct the assumptions that increased evidence will legitimise the field of sport-for-development, improve practice and enhance future policy. The authors own experiences, working as external evaluators, are also drawn upon to critique the value of current 'evidence'.

Findings

The chapter illustrates how current calls for evidence are somewhat misguided and are unlikely to fully realise the intended consequence of validating sport-for-development or improving future practice. Utilising personal reflections, the impact that Global North/Global South power imbalances have on data is highlighted, suggesting that this will rarely lead to data that provides a detailed understanding of work in practice.

Research limitations/implications

The chapter builds on the work of other authors highlighting the importance of disconnecting research from evidence and monitoring and evaluation in the sport-for-development field.

Originality/Value

The chapter utilises previous literature but also provides a rarely available personal perspective on the issue of evidence that continues to permeate the rationale behind undertaking research within sport-for-development.

Key words: Evidence, monitoring and evaluation, sport-for-development, impact

Introduction

Frequent calls for the development of an improved ‘evidence-base’ are ubiquitous in the field of sport-for-development. The following statements, taken respectively from a United Nations report, from a ‘Concept Paper’ developed by a Danish sport-for-development network and the opening editorial of the Journal of Sport-for-development demonstrate the widespread nature of such calls:

[Academia should] build a strong evidence base for the effective and efficient use of sport-for-development and peace that can feed into the development of viable policy recommendations. (United Nations, 2012, p. 21)

There is a growing pressure and interest to effectively measure and demonstrate the outcomes and impacts of sports projects. Monitoring and evaluation (M&E) are considered essential in the field, as it is vital for the process of continuous improvement and qualifications, and hence for the process of legitimising sport-for-development as a field. (Network for Sport and Development, 2009, p. 5)

SFD organisations must “evaluate or perish”. Only by applying rigorous research methods will the SFD sector establish adequate evidence to streamline its approach and survive broad contractions in foreign aid budgets. (Richards et al., 2013, p. 2)

We would argue that these quotes are representative of others by similar stakeholders in the field of sport-for-development in the urgency of their appeals and largely uncritical acceptance of the utility of evidence in influencing policy and practice. As can be identified in these quotes, calls for evidence are often associated with processes of monitoring and evaluation. These latter two processes of evidence-gathering are rarely distinguished,

although commonly they have a shared focus on a specific sport-for-development programme. Monitoring and evaluation can also, to differing extents, be considered as processes of research, although research may also encompass investigations that go beyond specific sport-for-development programmes.

Our purpose in this chapter is to problematise some aspects of the ‘call for evidence’ and the associated processes of monitoring, evaluation and research. In doing so, we add to some emergent critiques in the sport-for-development literature offered by Kay (2009; 2012), Nichols et al. (2011), Coalter (2013) and Adams and Harris (forthcoming) amongst others, whilst also recognising that such critiques are more common in other fields of development studies and public policy. The chapter is divided into two sections. The first of these sections examines and questions the rationales that underpin calls for evidence through considering the supposed benefits and beneficiaries of an improved evidence-base. The second section reflects more personally on these issues, critically discussing the realities of evaluation in practice through drawing on the experiences of the authors.

Evidence for What and Who?

As exemplified in the quotations above, there appear to be (at least) two assumed benefits of an improved sport-for-development evidence-base: firstly, more supportive policy and funding decisions and, secondly, improved practice and delivery of sport-for-development initiatives. Typically, it appears that the policy makers and funders who are to be convinced by improved sport-for-development evidence would belong to policy fields, such as international development, with a wider remit than solely sport-for-development. Conversely, the use of evidence to improve sport-for-development practice is an aspiration orientated towards stakeholders in the sport-for-development sector itself. There could be some connections between the two aspirations and groups of stakeholders, for example

through the assumption that improving practice will lead to more advantageous policy decisions. Nevertheless, this broad bifurcation of aspirations and stakeholders demonstrates that it is important to adopt a differentiated approach to examining the rationales underpinning calls for improved evidence of sport-for-development. This section will do so by considering the potential consequences of evidence on policy decisions and practice in turn.

If policy makers and funders are to be convinced of the merits of sport-for-development then it appears that the underlying assumption is that new evidence would indicate the positive developmental outcomes of sporting programmes. Thus, it would be easy to suggest that those making this argument would be amongst the much criticised sport-for-development ‘evangelists’ (Coakley, 2011; Coalter, 2013). Nevertheless, the view that the sport-for-development evidence-base is limited is not the preserve of such ‘evangelists’ alone. Even within some well-argued and critical academic literature, there can be claims that there is ‘little more than anecdotal evidence ... about the impact of sport in development’ (Hartman and Kwauk, 2011, p. 285). Nevertheless, as Coalter (2013) argues, this would be to ignore the substantial body of literature on the effects of sporting involvement and this literature includes comprehensive reviews by Coalter (2007) himself and the edited collection by Kidd and Donnelly (2007). To ignore such evidence completely appears to be associated with subscribing to the view of sport-for-development as a new ‘movement’ whilst ignoring the long history of sport being used for various social purposes (Kidd, 2008; Coalter, 2013).

It is possible to hypothetically consider the implications for the sport-of-development field if this existing, wider evidence-base were to be adopted by policy makers and funders, although we will argue later that considering the influence of evidence in such simplistic terms is unrealistic. A very brief summation of the wider literature would be that involvement

in sport ‘*might* lead to desired outcomes for *some* participants or some organisations in *certain circumstances*’ (Coalter, 2010, p. 311, emphasis as in original). Irrespective of any subsequent improvements in sport-for-development practice, it would be very unrealistic to believe that the rigorous generation of new evidence would contradict or alter this overall appraisal to any substantial degree. Therefore, apparent aspirations that sport-for-development as an entire field would be ‘legitimised’ (Network for Sport and Development, 2009) by collective evidence of positive impacts appear to be based on a misplaced optimism that is, in this case, characteristic of sport-for-development ‘evangelists’.

Rather, any adoption by policy makers and funders of rigorous evidence on sport-for-development outcomes would be more likely to have differential impacts for different stakeholders or particular programmes within the sport-for-development field. If existing evidence already qualifies, to some extent, what impacts are derived from particular types of involvement in sport, by which participants and in which contexts, then improvements to evidence on sport-for-development are likely to provide more fine-grained understanding of such qualifications. This evidence, then, would enable policy makers and funders to adopt a utilitarian prioritisation of particular programmes where impacts are considered to be more likely or more positive. For example, while we argue elsewhere of the inappropriateness of such measures (Lindsey and Culbertson, 2013), it is indicative to consider the potential consequences of the adoption of the evidence produced by Substance (2013) that sport programmes in the UK addressing substance misuse may result in broader cost savings that are over three times more valuable in financial terms per participant to those programmes addressing other social outcomes. No comparison with the estimated financial value of services other than sport on the same outcomes is offered. A hypothetically rational response to these findings would be direct a greater proportion of sport-for-development funding to

projects addressing mental health issues (and encouraging these projects to attract more participants), rather than necessarily increase the total funding to sport. Taken on a wider basis, such evidence-based policy and funding decisions would likely increase the divisions and competition between stakeholders in the already fragmented sport-for-development field. Pressure to produce positive results from evaluations of single programmes would also increase as a result, exacerbating issues described in the second section of this chapter.

Much of the preceding argument is based on policy makers and funders taking rational decisions based on available evidence. More broadly across different policy fields, ‘much of the debate around standards of evidence has focused on [such a] an instrumentalist view of evidence use, which involves the direct application of research to policy and practice decisions’ (Nutley et al., 2012, p. 14). This instrumental view appears to be the assumption adopted by those advocating for efforts to improve the sport-for-development evidence-base, and certainly there are few voices within the sport-for-development field speaking to a more nuanced view. Elsewhere, however, Nutley et al. (2012) is joined by others such as Mulgan (2005) and Smith and Joyce (2012, p. 57) who identify that

such a linear relationship [between evidence and policy] is dependent on an unrealistically simple account of policy making, given that policy decisions are also informed by a multitude of other factors, including ideology and values, public opinion and lobbying.

The veracity of this criticism becomes further apparent if we consider theoretical models of the policy making process. Whilst rational processes in which decisions would linearly derive from consideration of the evidence are still promoted as appropriate approaches *for* policy development (e.g. Veal, 2002), few contemporary policy analysts would argue that such models represent an accurate model *of* the policy process in actuality (Hill, 2005). Instead,

Houlihan (2005) argues for the utility of models, such as Multiple Streams and Advocacy Coalitions Frameworks, which offer more complex explanations of the policy process. Within the Multiple Streams Framework, evidence may help to identify particular policy problems or support solutions to these problems. However, this framework also highlights the unpredictability of the policy process in which individual ‘policy entrepreneurs’ play a significant role in enabling particular policies to be adopted. For Smith & Joyce (2012, p. 66), the Multiple Streams framework

illustrates the unpredictability of knowledge transfer and suggests that there will always be limitations to guidance on how to promote the increased uptake of research evidence.

Technical information, of which the types of evidence discussed in this chapter are but one form is considered to have a more specific role in the policy process within the Advocacy Coalition Framework. Nevertheless, policy remains substantially determined by the interaction of different coalitions of actors each with their own set of beliefs which may be (or equally may not be) adapted according to evidence and only over a relatively long-period of time (Houlihan, 2005).

The point of considering such models, which serve largely as examples in this case, is not to fully identify the policy making process as it affects sport-for-development. Rather, consideration of each of these models, and others, helps identify the limited influence that evidence may have on policy making. Moreover, the models suggest that evidence, of the rigour being called for, would be but one form of information that may be utilised by those involved in the policy process. Considered alongside our earlier qualifications as to what evidence may actually indicate about the impact of sport-for-development initiatives, it becomes clear that the aspirations of those who may identify evidence as a ‘game changer’ in

establishing sport-for-development within broader but related policy fields are ultimately likely to be disappointed. As we turn to some of the issues related to the potential contribution of further evidence to sport-for-development practice, it becomes apparent that similar misconceptions occur.

If evidence regarding sport-for-development is to influence practice then it is important to initially consider the nature of the evidence that would be useful in doing so. Such a consideration is often not present in undifferentiated calls, such as those cited in the introduction, to extend the sport-for-development evidence-base. Nevertheless, a relatively self-evident initial observation would be that the types of information that may be beneficial to improving practice within the sport-for-development sector would not likely be the same as those that would (supposedly) convince external policy makers and funders of the merits of sport-for-development¹. Such a distinction broadly corresponds to Coalter's (2013) argument that, rather than concentrating on demonstrating impact, the sport-for-development sector needs to develop greater understanding of the processes or mechanisms that may lead to such impact. This is not to say that different types of information could, and in some cases are, being collected through the same evidence-gathering process. However, it is important to recognise that practices of monitoring, evaluation and research may each focus to a different extent on sport-for-development processes or outcomes, and, therefore, may each have different degrees of utility for improving sport-for-development practices.

By their very nature, practices of monitoring and evaluation are concerned with specific, and mostly singular, sport-for-development programmes. It is these specific sport-

¹ This is not to say that improving sport-for-development practice may not contribute to efforts to garner support amongst external policy makers but, as the arguments early in the section indicate, this is very unlikely to be a linear or short-term process.

for-development programmes that are most likely to be improved as a result of being monitored or evaluated. However, a significant proportion of such evidence gathering in sport-for-development is undertaken by external consultants at the behest of programme funders. While the implications of these approaches to evidence gathering will be considered in the following section, it is somewhat inevitable that subsequent reporting and dissemination practices will be orientated towards those organisations that have commissioned and funded monitoring, evaluation or research (Adams & Harris, forthcoming). The extent to which reported findings are valuable to directly inform the practice of those delivering programmes must, therefore, be questioned, irrespective of the extent to which funding organisations may attempt to influence subsequent implementation practices.

What is even more doubtful is the extent to which monitoring, evaluation and potentially research may contribute to the improvement of practices enacted in the delivery of sport-for-development programmes on a wider basis. There are, at least, three initial reasons to doubt whether this may occur. First, as Cronin (2011) recognises, weaknesses of dissemination within the sport-for-development sector limit the utility of relevant evidence. For example, a substantial proportion of programme evaluations are not made public and academic research is often published in formats that are both directly and indirectly inaccessible to practitioners. Second, Cronin (2011) identifies that existing reviews of sport-for-development evidence tend to focus on impact rather than underlying processes. In other international development fields, systematic reviews have attempted to draw broader lessons from drawing together evidence from different programmes (DfID, no date). Further methodologically rigorous reviews, focused on practices (rather than impact) and aligned with specific themes of sport-for-development work, may be considered beneficial for

improving future practice. However, it is our third concern that is perhaps the most fundamental. Irrespective of any progress made with regard to the synthesis of evidence or its dissemination, any suggestion that findings from particular sport-for-development programmes could be beneficial to practice in other programmes could be considered tenuous. The reason for this claim is that the contexts within which sport-for-development programmes are delivered, and their intended beneficiaries, may differ in infinitely different and unknown ways. The effects of these potential differences on sport-for-development processes and outcomes certainly cannot be identified directly from the evidence pertaining to the initial programmes. As a consequence, we must consider that transferability or generalisability of evidence-based practices may be inherently limited even if there is comprehensive evidence regarding the relationship between specific programmes and the context in which they are delivered.

Reflecting on monitoring and evaluation practices

As discussed in the introduction, the second half of the chapter will consider some of the broader arguments presented in the previous section and explore some of the realities of attempting to collect a 'strong evidence-base' in practice. We will do so by drawing on our experiences of evaluating several sport-for-development projects delivered by Non-Governmental Organisations (NGOs) in African cities over a period of seven years. Whilst different international development agencies provided the funding for these evaluations, UK-based sports agencies were responsible for the day to day overseeing of in-country stakeholders and delivers. Five different NGOs were involved in the evaluated projects, two of which were indigenous, developed and managed by local people, the other three were international NGOs but managed by locals in-country. This section builds on the previous by also questioning the possibility and value of collecting 'robust' evidence and highlighting

some of the challenges personally experienced when attempting to do this. The first part discusses how Global North/South power dynamics influence the process of monitoring and evaluation, including how this shapes the relationship between evaluators and NGO staff, how monitoring and evaluation is perceived by NGO staff and in turn how this affects the type of data external evaluators are able to collect. The second part considers how Global North donors may shape the type of data that is collected and how relevant and valuable this actually is for NGO staff working on the ground.

The discussion that follows reflects explicitly on donor funded projects and donor funded external evaluation systems undertaken for the ‘internal audiences’ of the UK sporting and development agencies that managed and funded the projects respectively. We do recognise that other models of evaluation take place including locally initiated evaluation and donor imposed self-evaluation by NGO staff. Whilst the sporting agencies had some interest in understanding delivery mechanisms, the interests of the development agencies were primarily centred on gaining evidence of impact and justification of future budgets. These complexities contribute further to the points developed in the previous section by both demonstrating the overlapping purposes of specific evaluations and indicating that evaluations may be commissioned for somewhat narrow self-interest amongst organisations already committed to funding sport-for-development. Grand aspirations of ‘legitimising sport-for-development as a field’ (Network for Sport and Development, 2009, p. 5) appeared absent.

From our experiences of working on this type of evaluation, we would argue that requests for evidence and the belief that robust evidence can be collected seemingly ignores the inherent power dynamics that exist between Global North funding agencies, Global North researchers, Global South NGOs and on-the-ground deliverers. These dynamics, as we will

discuss, affect the entire evaluation process continually shaping the type of information that can be collected. The imbalances of power that exist between donor funding agencies and NGOs have been well established within international development literature (Easterley, 2002; Ebrahim, 2003) and there is a growing recognition within sport-for-development of how these shape and influence practice (Beacom, 2007; Darnell and Hayhurst, 2011; Kay, 2012). The hierarchical lines of power that are transmitted down from funding agencies to NGOs inevitably exert considerable influence on how evaluation is interpreted and understood by in-country staff and participants in communities. Supporting the points we raised earlier in the chapter, whilst evaluation is often ‘sold’ to NGO staff as supporting and enhancing their practice, the NGOs we worked with rarely saw it as anything other than an assessment of their capabilities and capacity to deliver impacts determined by funding agencies. From the outset therefore, the ability of us as external evaluations to capture the day to day realities of what happens within sport-for-development projects in any sort of meaningful or useful manner is limited. For NGOs we worked with, concerned about further funding, providing anything other than a positive picture of their practice would be a foolish and risky endeavour. Inevitably, what is presented to external evaluators tends to be ‘packaged’ by NGO staff to provide a particular image that they believe the funding agency wants to see.

It is useful to illustrate what NGO staff believe is at stake when engaging with evaluation via a specific example. One project worked on provided funding for the employment of additional NGO staff to oversee the program across several different provinces. These positions were filled by individuals who had previously been working in the capacity of voluntary peer leaders on other projects with the NGO. They were from highly impoverished compound communities. For these individuals therefore, the UK funded roles

represented a significant opportunity, providing regular paid employment that generally would have been unavailable to them within their compound communities. They had previously relied on piecemeal, unsecured paid work for their livelihoods and voluntary opportunities to develop their skills. The positions with the NGO provided a level of security for both them and their families that simply would not have been experienced previously. This context is important because it emphasises what was at stake for these staff members when engaging with evaluation. For them, as much as we may try and suggest evaluation is undertaken to facilitate learning and program development, its connections with accountability, assessment and ultimately judgements that may have significant consequences for their livelihood were inescapable. Inevitably, staff became very nervous of communicating information that may suggest they were not meeting the ambitious outcomes of the funding agency and this has considerable implications for how they will assist external evaluators.

‘Outsider’ evaluation experts are considered to best placed to provide the type of rigorous, objective review of practice that is required to produce the ‘holy grail’ of a robust evidence-base (Riddell, 1999). However, our experiences would suggest it is naive to assume that external evaluators work in a social vacuum. Instead evaluators become an intricate aspect of the top-down power relations between Global North funders and Global South recipients (Kay, 2012). Evaluators become the interface between the programs on the ground and how external donor agencies view and understand them. Within other studies (Jeanes & Kay, 2013), we have provided a fuller analysis of the inherent tensions and possible repressive power relationships that exist between Global North researchers contracted to undertake evaluation and Global South NGO staff and communities. These relations greatly restrict the type of data and information that can be collected. The context in which

evaluation data is usually collected also does little to assist with reducing the North/South hierarchy. Within sport-for-development, UK funders have, at least, tended to prefer the use of UK-based researchers. There are possibly a number of reasons for this, a perceived lack of capacity within country, potentially a lack of trust of the objectivity and rigour of in-country assessments and conversely a belief that UK based researchers have greater knowledge and capacity to evaluate in a way that is considered trustworthy and reliable. With limited finance available, the practicalities of this approach means that evaluators such as ourselves tend to spend, short, intensive periods of time in Global South countries. This has further impacts on the type of data collected. Practically, as an evaluator to collect any data it was essential to develop a strong working partnership with the NGO staff. Ethically this is also desirable particularly if seeking to attempt to develop a more collaborative approach to evaluation that gives, at least, some power back to the NGO. We found ourselves as external evaluators often entering contexts with limited local knowledge and virtually no established on the ground relationships, the support of NGO staff was absolutely essential to gain both access and cooperation from project participants to collect data. Consequently, the idea that we can somehow be more objective and rigorous becomes questionable. Inevitably we visited communities that NGO staff felt would provide the best image of the project in action, NGOs presented young people who they were confident would communicate positive experiences and, importantly, who would demonstrate to funding agencies that their money was being well spent.

These tensions were perhaps most apparent when evaluating multiple projects simultaneously, all of which were funded by different agencies and run by different NGOs in the same area. One young girl appeared at several different interviews providing a fulsome and rich account of how each project had been invaluable to her and her community. She was

extremely articulate and, without prompting, discussed a complex and difficult childhood, explaining how each sport development project had provided her with skills, hope and a focus to develop herself. Staff at different NGOs had clearly recognised that she would provide a compelling description of how much she valued each project. As this young girl was used as a figurehead across several projects, it prompted a consideration by us of whether she was representative of the broader community of young people taking part in the projects. Generally though, accessing a random and diverse selection of participants is almost impossible because of the reliance on NGO staff to facilitate data collection.

It is practicalities such as these described above that we must consider when seeking to answer calls for more 'evidence'. As outsiders, whose reports are perceived to stimulate further funding or result in a cut to existing funding, evaluators are in an extreme position of power and authority. Inevitably, NGO staff respond by restricting the data external evaluators can access and ensure what evaluators can 'reach' presents a positive image of the value and impact of the work being undertaken. Our own experiences highlight how monitoring and evaluation can actually provide a context where NGOs resist power hierarchies and exert some agency by controlling the information that is passed back up the power chain. To consider that they would do otherwise is naïve and greatly under-estimates the authority external evaluators hold. This discussion therefore raises considerable questions as to whether objective, robust and accurate data can be collected within an evaluation framework. Given the constraints and complexities, the likely answer is no and the continued call and quest for this type of knowledge fails to acknowledge or grossly oversimplifies the relationships and the repressive power dynamics that exist between funding agencies, NGOs, evaluators and participants.

As external evaluators we have been more fortunate than many in that we have been able to work with the same NGOs across a number of years in evaluating different projects. This ongoing opportunity has allowed the establishment of deeper and potentially more trusting relationships with NGO staff as the years have progressed. As reports have been developed in consultation with NGO staff and well received by funding agencies we have potentially become to be seen as more of a help and less of a hindrance. A consequence of this has been the opportunity to talk with staff in a more relaxed and informal manner. Additionally, we have gained a much better understanding of the intricacies of the communities they are working within. Through this, we have been able to speak with a broader range of young people, see more of the day to day realities of projects in action and recognise that, although it may feel that we are being engineered towards such a conclusion, many young people do gain considerable benefit from their involvement (Jeanes, 2013). When returning to communities year after year and meeting the same peer leaders, they also have become more open about the challenges and difficulties they face. Often they would talk about the difficulties of persuading young people to change their behaviour when there is little support from broader family. They admit that there are frequently large numbers of young people within their community who sport-for-development programmes simply cannot reach and work with.

Similar to Spaaij (2011), we have concluded that impacts are likely to be limited by the broader socio-political context in which projects are operating and therefore achieving sustainable impact is inherently challenging. We have also recognised that 'absolute' understanding or 'robust' evidence along Global North lines simply is not possible. We have gained greater insights and understanding from these more fulsome interactions but the additional knowledge continues to create more uncertainties and ethical tensions. The more

detailed information that has begun to emerge in informal discussions is highly valuable for understanding why particular approaches may be problematic or limited but we have had concerns that sharing such information with funding agencies would again potentially lead to negative judgements on the viability and impact of particular projects. Such is the competitive climate of funding for sport-for-development projects, there always remains the concern for NGO staff that highlighting particular weaknesses with a project will have a detrimental effect on project funding, even when these limitations are the result of broader structural issues beyond the control of project workers.

We have questioned earlier in the chapter who evidence is being collected for and for what purpose. As the previous section suggests, formalised systems do not necessarily allow the type of data to be collected that NGO staff would potentially find useful. Alongside generic calls for extending the evidence-base of sport-for-development, practitioners and academics have advocated for more standardised, theory driven approaches to evaluation (Coalter, 2007, 2010; Schulenkorf, 2012). Approaches to evaluation based on 'logic models' have been increasingly popular within sport-for-development to attempt to provide consistency and transparency on what projects are attempting to achieve and how it is believed that this will happen (Savaya & Waysman, 2005). Within broader development studies, there has been a range of critiques of the inflexibility of logic model approaches and the tendency for them to again lead to top-down program delivery and evaluation. Chambers and Pettit (2004) have critiqued how models tend to largely be devised by funding agencies that determine the aims and intended outcomes of particular projects and then pass these on to local staff to implement. This was reflective of our experience across several projects. During initial discussions with one UK sporting agency about a potential evaluation, the UK based lead staff member commented 'the logic model we have for the project is absolutely

ridiculous. It's talking about working with thousands of year people and achieving all these things that are impossible..... we know it is ridiculous but it's what [the funding agency] want to see happen'. Logic models for many of the projects worked on have been constructed by funding agencies in isolation and handed back to on the ground staff to implement, as literature suggests has previously occurred elsewhere (Chambers & Pettit, 2004; Hinton, 2004).

We have also worked on projects where logic models have been constructed in collaboration between funding agencies and NGO staff. The experiences of this process have been similar to those documented by Win (2004) who describes her experiences of trying to develop a logic model under the guidance of foreign aid staff,

We spent three days trying to fit visions, objectives, strategies and our way of seeing the world into the differently shaped blue, green and yellow cards. It was really not funny, though. It was painful. Nobody understood this method and the logic behind it (Win, 2004, p. 125).

Similarly the NGO staff we worked with struggled to see the relevance of the logic model, were unsure what was required of them to 'populate' the model and ended up deferring to Global North's staff opinion on what their logic model should look like just to 'get it done'. The potential for the resultant logic model to continue to influence the practice of these NGO staff for the subsequent years has not been appreciated in this process. Whilst on the surface the process is seemingly undertaken in partnership, in reality the outcomes that NGO staff are working towards continue to be prescribed to them. Kay (2012, p.892) has similarly highlighted 'although M + E ostensibly serves to inform programme development, in reality it is inevitably shaped by the organizational requirements of those commissioning it'. This often leads to project aims that are both unrealistic and often do not reflect local need. Guest

(2009) has highlighted that outcomes attributed to sport-for-development projects tend to focus on constructs established within the Global North such as self-esteem, that have limited relevance or meaning in the Global South. Guest discusses how such targets are arbitrary and generally meaningless within the context of project communities, an issue we also experienced. The logic models guiding the evaluations we have worked on have required NGO staff to meet targets of raising self-esteem levels across particular numbers of participants. They have also provided quantitative targets for number of young people who are empowered, a concept in itself which is incredibly complex and difficult to measure. Funding agencies tend, via their logic models, to place emphasis on numbers at the expense of understanding the experiences that occur within programs. One particular project set NGO staff a target of training approximately 100 peer leaders over the period of a year. Naturally, staff felt they had to do this even if such numbers were not required to facilitate delivery of sport-for-development activities in communities. The emphasis shifted from an informal system of training that was previously in place to one where getting anybody to undertake formal training became paramount so that staff could 'meet their target'. Peer leaders who had been organising delivery for many years had attended training to help increase numbers but also many young people were being trained who were not necessarily returning to communities and subsequently delivering sport-for-development activities. Whilst it could be argued that they may have gained some valuable skills via the training, conversely the emphasis on numbers was stretching the NGOs limited resources and preventing staff from working with active peer leaders on a more regular basis. Instead staff time was largely taken up by organising numerous training sessions across communities and persuading young people to attend. Again, the needs of the NGO and especially the local communities had become lost amongst what funding agencies thought was needed. Because the collection of

evidence becomes tied to ‘proving’ such targets have been met, local nuances become lost within the monitoring and evaluation process.

Conclusions

This chapter has attempted to problematize the extensive and increasing calls for improvements in the evidence-base of sport-for-development programs. It has done this by raising questions across a variety of dimensions that consider, what evidence is required, for whom, to serve what purpose and how this evidence is collected in practice. Through analysis of these questions we have pinpointed a range tensions that arise in the quest for evidence. The first half of the chapter in particular has highlighted the lack of specificity that exists in current calls for developing an evidence-base. These currently generally fail to consider what type of evidence is needed and for what purpose and, by implication therefore, suggest that any type of evidence is a positive thing and will enhance the credibility of the sport-for-development field. As the first half of the chapter points to, such beliefs are contentious.

Several assumptions have tended to drive the focus on collecting evidence, that it will ‘prove’ the value of sport-for-development, it will lead to enhancement of future practice and it will guide the development of future policy. All of these beliefs have been deconstructed and illustrated to be questionable. Firstly, we cannot assume that such evidence if it is collected will provide a positive assessment of sport-for-development projects, secondly, broader policy analysis has suggested evidence is rarely at the forefront of guiding future policy developments and, finally, we are unsure whether current ‘evidence’ is of practical value to NGO staff or whether it even reaches them. The second half of the chapter has highlighted how external monitoring and evaluation, as a specific tool through which to collect evidence, is inherently problematic. We have suggested that Global North/Global South power imbalances will continually impact the type of evidence that is collected via

external monitoring and evaluation processes, leading to questions about the value of this information particularly when intending to use such knowledge to guide future policy and enhance practice. Such information instead tends to satisfy funding agency requirements and potentially lead to future funding opportunities but does little to provide enhancements to the field that it is suggested evidence will provide.

As Kay, (2012) and Coalter, (2010) advocate, there needs to be a move away from the notion of collecting evidence towards one that prioritises the development of understanding. Kay (2012, p. 900) highlights that in order to achieve this there is a need to disentangle monitoring and evaluation from its connections with funding and ultimately 'evidence'.

There is an argument, therefore, for reorientating M+E from the interminable pursuit of 'definitive' evidence, which primarily addresses the priorities of external funders, to a quest for alternative types of knowledge that may prove more appropriate, valid and obtainable and offer more value to sport in development programmes in-country.

We would additionally suggest a distinction should occur between research and monitoring and evaluation more broadly. Whilst the three may be interchangeable to in country staff, we would also argue that the differences also are not readily recognised within Global North-led calls for evidence. However we believe a research focused approach would move away from examining specific sport-for-development projects in isolation and allows opportunities to develop knowledge of wider social contexts and processes to develop the type of understanding Kay (2012) and Coalter (2010) suggest would be valuable and also support improvements in practice. The work of Spaaij (2011) in Brazil is one such example of this type of detailed research. Kay (2012) again makes a valuable point on this issue suggesting that knowledge that local staff and deliverers may find useful does not necessarily have to be

generated through formalised research channels. She highlights how in projects she researched in Brazil and India,

Knowledge and learning about how to use sport came not from a structured system of M + E....: It was underpinned instead by detailed understanding developed through many years spent in local schools and communities. The projects were successful and dynamic and project learning was taking place....but formal information gathering systems played little role in it (Kay, 2012, p. 897).

This highlights an important issue, that knowledge and understanding is likely to be generated spontaneously within the projects themselves and does not necessarily have to come from external researchers. It is important to recognise the potential role we, as researchers, also have in assisting NGO staff and deliverers with sharing such knowledge.

In the opening section of the chapter we have highlighted that the calls for evidence within sport-for-development are focused upon using such information to provide broader legitimisation of the field. We would suggest that the type of research we advocate, which shifts away from evidence and in particular validating notions of the ‘power of sport’, may actually assist in the longer term with this aim. More nuanced and subtle understandings are likely to enhance practice and contribute to recognition of sport-for-development as a more mature field that can make a contribution to a broader development effort. The challenge is for those involved in the gathering of evidence to modify their practices so that they can make a more effective contribution to advancing the sport-for-development field in this way in the future.

Five Key Reading:

- 1. Spaaij, R. (2011) *Sport and Social Mobility: Crossing Boundaries*. New York and London: Routledge.**

Ramon Spaaij's book and work more generally provides excellent examples of the value of the detailed, localised research that we suggest is necessary to develop the field of sport-for-development.

- 2. Lindsey, I. and Gratton, A. (2012) An 'international movement'? Decentering sport-for-development within Zambian communities, *International Journal of Sport Policy and Politics*, 4:1, 91-110.**

This paper also provides a useful illustration of the type of information that can be gathered by ethnographic studies that move away from understanding single projects in isolation and instead seek to develop whole community knowledge. The paper also highlights the potential contradictions that exist between evidence based knowledge and data collected in this more nuanced way.

- 3. Kay, T. (2012) Accounting for legacy: Monitoring and evaluation in sport in development relationships, *Sport in Society*, 16 (6): 888-904**

Tess Kay's paper offers an insightful critique of the relationship between monitoring and evaluation, funding agencies and accountability within sport for development. The current chapter seeks to build on several of the core arguments presented in this paper.

- 4. Jeanes, R. & Kay, T. (2013) Conducting research with young people in the global south, in K te Riele & R Brooks (Eds) *Negotiating Ethical Challenges in Youth Research*, London: Routledge, pp. 19-30.**

This chapter provides greater insights into some of the issue raised within the current study regarding imbalances of power that exist between Global North researchers and Global South NGO staff and sport-for-development participants and how this impacts on data collection.

5. Coalter, F. (2010) The politics of sport-for-development: limited focus programmes and broad gauge problems? *International Review for the Sociology of Sport*, 45: 295-314

Fred Coalter's work generally and this paper specifically is useful for understanding the rationale underpinning monitoring and evaluation within sport-for-development and also problematising its current value.

References

- Adams, A. & Harris, K. (forthcoming) 'Making sense of the lack of evidence discourse, power and knowledge in the field of sport-for-development' *International Journal of Public Sector Management*.
- Beacom, A. (2007) A question of motives: Sport reciprocity and international development assistance. *European Sports Management Quarterly*. 7 (1): 81-107.
- Chambers, R. & Pettit, J. (2004) Shifting power to make a difference. In L. Groves & R. Hinton (Eds) *Inclusive Aid: Changing Power and Relationships in International Development*, London: Earthscan, pp.137-162
- Coakley, J. (2011) 'Youth Sports: What Counts as "Positive Development?"' *Journal of Sport & Social Issues*. 35 (3) pp. 306-324.
- Coalter, F. (2007) *A Wider Social Role for Sport: Who's Keeping the Score?* Oxon, UK: Routledge
- Coalter, F. (2010) The politics of sport-for-development: limited focus programmes and broad gauge problems? *International Review for the Sociology of Sport*, 45: 295-314
- Coalter, F. (2013) *Sport-for-development: What Game are we Playing?* Abingdon: Routledge.
- Cronin, O. (2011) *Comic Relief Review: Mapping the Research on the Impact of Sport and Development interventions*. Manchester: Orla Cronin Research.
- Darnell, S. and Hayhurst, L. (2011) Sport for decolonization: Exploring a new praxis of sport-for-development. *Progress in Development Studies*, 11 (3), 183- 96
- Department for International Development (DfID) (no date) Research for Development: Systematic Reviews. <http://r4d.dfid.gov.uk/SystematicReviews.aspx>. [accessed 11/09/2013].

- Easterley, W. (2002) The cartel of good intentions: The problem of bureaucracy in foreign aid, *The Journal of Policy Reform*, 5 (4): 223-250,
- Ebrahim, A. (2003) Accountability in practice: Mechanisms for NGOs. *World Development* 31, (5): 813–829
- Guest, A. (2009) The diffusion of development-through-sport: analysing the history and practice of the Olympic movement's grassroots outreach to Africa. *Sport in Society*, 12 (10), 1336-1352
- Hartmann, D. & Kwauk, C. (2011) 'Sport and Development: An Overview, Critique, and Reconstruction' *Journal of Sport & Social Issues*. 35 (3) pp. 284-305.
- Hill, M. (2005) *The Public Policy Process*. 4th ed. Harlow: Pearson Education.
- Hinton, R. (2004) Enabling inclusive aid: Changing power and relationships in international development. In L. Groves & R. Hinton (Eds) *Inclusive Aid: Changing Power and Relationships in International Development*, London: Earthscan, pp. 210-220
- Jeanes, R. (2013) Educating through sport? Examining HIV/AIDS education and sport-for-development through the perspectives of Zambian young people, *Sport, Education and Society*, 18 (3): 388-406
- Jeanes, R. & Kay, T. (2013) Conducting research with young people in the global south, in K te Riele & R Brooks (Eds) *Negotiating Ethical Challenges in Youth Research*, London: Routledge, pp. 19-30.
- Kay, T. (2009) 'Developing through sport: evidencing sport impacts on young people' *Sport in Society*. 12 (9) pp. 1177-1191.
- Kay, T. (2012) Accounting for legacy: Monitoring and evaluation in sport in development relationships, *Sport in Society*, 16 (6): 888-904

- Kidd, B. & Donnelly, P. (eds) (2007) *Literature Reviews on Sport-for-development and Peace*. Toronto: University of Toronto.
- Kidd, B. (2008) 'A new social movement: Sport-for-development and peace' *Sport in Society*. 11 (4) pp. 370-380.
- Lindsey, I. & Culbertson, L. (2013) The holy grail of evidence-based policy ... and its distortion of sport-for-development practice. *European Sports Development Network Conference*. Edge Hill University. 4th-5th September 2013.
- Mulgan, G. (2005) 'Government, knowledge and the business of policy making: the potential and limits of evidence-based policy' *Evidence & Policy: A Journal of Research, Debate and Practice*. 1 (2) pp. 215-226.
- Nicholls, S., Giles, A. R. & Sethna, C. (2011) 'Perpetuating the 'lack of evidence' discourse in sport-for-development: Privileged voices, unheard stories and subjugated knowledge' *International Review for the Sociology of Sport*. 46 (3) pp. 249-264.
- Nutley, S., Powell, A. & Davies, H. (2012) 'What counts as good evidence?' <http://www.nesta.org.uk/library/documents/A4UEprovocationpaper2.pdf>; [accessed 11/09/2013].
- Richards, J., Kaufman, Z., Schulenkorf, N., Wolff, E., Gannett, K., Siefken, K. & Rodriguexz, G. (2013) 'Advancing the Evidence Base of Sport-for-development: A New Open-Access, Peer-Reviewed Journal' *Journal of Sport-for-development*. 1 (1) pp. 1-3.
- Riddel, R. (1999) Evaluating NGO development interventions in D. Lewis (Ed.), *International Perspectives on Voluntary Action: Reshaping the Third Sector*, London: Earthscan, pp. 222–241

- Savaya, R. & Waysman, M. (2005). The logic model: A tool for incorporating theory in development and evaluation of programs. *Administration in Social Work*, 29, 85-103.
- Schulenkorf, N. (2012) Sustainable community development through sport and events: A conceptual framework for Sport-for-Development projects, *Sport Management Review*, 15 (1): 1-12.
- Smith, K. E. & Joyce, K. E. (2012) 'Capturing complex realities: understanding efforts to achieve evidence-based policy and practice in public health' *Evidence & Policy: A Journal of Research, Debate and Practice*. 8 (1) pp. 57-78.
- Spaaij, R. (2011) *Sport and Social Mobility: Crossing Boundaries*. New York and London: Routledge.
- Substance. (2013) 'Sports works' <http://www.sported.org.uk/what-we-do/sportworks>; [accessed 11/09/2013].
- United Nations. (2012) 'Sport-for-development and peace: mainstreaming a versatile instrument.' http://www.un.org/wcm/webdav/site/sport/shared/sport/pdfs/SG%27s%20Reports%20to%20GA/A-67-282/A-67-282_English.pdf . [accessed 11/09/2013].
- Veal, A. J. (2002) *Leisure and tourism policy and planning*. 2nd ed. Wallingford: Cabi Publishing.
- Win, E. (2004) If it doesn't fit on the blue square it's out! An open letter to my donor friend. In L. Groves & R. Hinton (Eds) *Inclusive Aid: Changing Power and Relationships in International Development*, London: Earthscan, pp.123-7

