11   Researching and Evaluating sport development and sport for development

Emma Sherry Nico Schulenkorf Pamm Phillips Katie Rowe

INTRODUCTION
To ensure that sport development (SD) and sport for development (SFD) programs achieve their intended outcomes, effective monitoring, evaluation, and learning (MEL) is crucial. Evaluation of both sport and non-sport outcomes ensures that the organisational mission and strategic goals are adhered to and that the program is delivered in the intended manner, at the intended time and with the intended participants (Coalter 2013a). Developing a clear and informed understanding of the success – or failure – of a program, event, or initiative in achieving its intended outcomes provides vital information for sport managers, coaches, and decision makers. By analysing why a program was successful or not, lessons can be learned about the context, the approach, the tools, and the intent of the program, which can also facilitate the design of future initiatives. The planning, management and delivery of sport programs represent considerable investments in both time and money, and it is imperative that organisations regularly monitor and evaluate their progress and outcomes. In other words, MEL ensures that resources and funding are being well spent (Houlihan 2011).

Organisations that provide SD or SFD programs require funding to design, create and implement them. Funding organisations may be a government body or a major state, national or international sporting organisation, such as the IOC. Sometimes, program funding will be provided by commercial partners or through a collaboration of private philanthropic and government agencies; for example, the Grootbos Foundation is a conservation organisation that teamed up with Barclays and ABSA banks and the Western Cape Department of Cultural Affairs and Sport to create the Football Foundation of South Africa (FFSA), an organisation delivering programs aimed at providing a sporting infrastructure legacy from the 2010 FIFA World Cup, held in South Africa (Swart *et al*. 2011).

In a sporting context, competition for funding is generally high (Houlihan 2011), meaning that funding organisations are required to prove that they have directed and used funds appropriately (Coalter 2011). Each external funding organisation has its own strategic goals and objectives; therefore, sports organisations seeking to acquire funds through these sources need to demonstrate how their program meets those strategic targets. Regular monitoring and evaluation of the programs is key to examining and highlighting how these objectives have been met or addressed, and it creates learnings for future planning and delivery. At the same time, professional monitoring and evaluation can demonstrate the proper use and acquittal of funds which, in turn, may help to secure ongoing commitment and additional financial support.

OUTLINE OF THE CHAPTER
This chapter discusses the role of monitoring, evaluation and learning in SD and SFD programs. It introduces and contextualises the challenges and opportunities for assessing programs and provides an overview of evaluative tools, approaches, and theory.
After completing this chapter, you should be able to:
   a. Provide an overview of SD and SFD evaluation approaches and challenges.
   b. Understand the link between stakeholders, funders and evaluation.
   c. Assess evidence-based policy and programming.
   d. Distinguish between research and evaluation.
   e. Engage with SD and SFD evaluation approaches and challenges.
   f. Critically discuss the role of the researcher and research teams ("Combining knowledges").

WHAT DOES THEORY TELL US?
Evaluation of outcomes begins with program design (Coalter 2013a). Research in the SD field has demonstrated that while this sounds like a simple enough statement, it is in fact more complicated than it first appears (Houlihan 2011). In particular, managers of programs must identify the specific outcomes they desire from the program and, simultaneously, consider how they will measure the success of those outcomes (Coalter 2013a; Long 2011).

The example of an SFD program aimed at reducing crime in a disadvantaged neighbourhood can demonstrate the complexity of this task. Many factors influence how such a program might be designed. One of the first considerations is the participants. Who is the program aimed at? How has this cohort of participants been selected? The identification of

target participants is key to ensuring that the desired impacts are achieved; in our example, it would be necessary to understand who is committing the crime in the neighbourhood by gathering evidence. At the same time, the "why" question is critical and information about the neighbourhood itself – the social context – is important to make informed decisions.

A second consideration is to understand the intended outcomes. In other words, what – specifically – is the program hoping to achieve? Crime reduction is a desired overall outcome, but how will the program achieve that end? Will it be through improving individuals' well-being with the intent that improved self-esteem will lead to a reduced desire to commit crimes? Perhaps it will be through providing a place for disaffected people to spend time, keeping them occupied and involved in a social setting in which they can make friends and social connections, which will in turn reduce the desire to engage in antisocial behaviours? Perhaps it will be by providing a place to learn new skills and develop talents that could open new pathways?

If the aim of our SFD program is to improve well-being, key questions from an evaluation perspective include:
- How will well-being be measured and defined? How will improvements in well-being be measured?
- How will the activities that are likely to lead to improved well-being be determined? Given the subjective nature of well-being, how is a program that impacts positively on a cohort of individuals' well-being designed appropriately?
- Is the chosen determination of well-being commensurate with that of our intended participants? In other words, might our own personal understanding of the necessary conditions for well-being differ from those of intended participants?

Similarly, an SD program might seek to grow the sport into specific regions or cohorts. If this is the case, key questions from an evaluation perspective might include:
- How might the intended groups be identified?
- What are the specific conditions surrounding these cohorts' non-participation? Are there financial constraints – such as residing in a socioeconomically disadvantaged area?
- Are there locational or geographic constraints – such as remoteness?
- Perhaps the sport is broadly considered a masculine sport, meaning females struggle to participate?
- Are there cultural constraints? Perhaps the sport is not well known in a particular cultural group; perhaps in that culture, sport participation is considered inappropriate for one gender?
- Are there personal physical constraints, such as a disability?

The intended outcomes must also be considered: does the program seek to develop individual athletes or teams (SD focus) or is it aimed at raising overall participation and engagement in sport and physical activity (SFD focus)? The answers to questions such as these can facilitate program design in that they can act as a guide for managers to ensure that program aims for outcomes are evidence-based as well as Specific Measurable, Achievable, Repeatable and Timely – often referred to as SMART goals.

PROGRAM THEORY

Program theory offers a framework through which to understand the relationship between program design and program outcomes. Adoption of a theory-based approach to sport programming enables sound evaluation and closes the gap between policymakers, practitioners, and researchers by providing opportunities for collaboration and conversation (Coalter 2013a). Through discussion, the finer details of program intents are brought into the light, allowing assumptions to be examined and a clearer picture of the issues, the situation, and the desired outcomes, which facilitates and supports evaluation (Coalter 2013a). The use of a program logic model (Figure 11.1) can be readily applied to both the SFD and SD settings. Coalter (2013a) discusses the theory in terms of *inputs*, *outputs*, *impacts* and *outcomes*, as Figure 11.1 demonstrates the application of this in an SFD context.
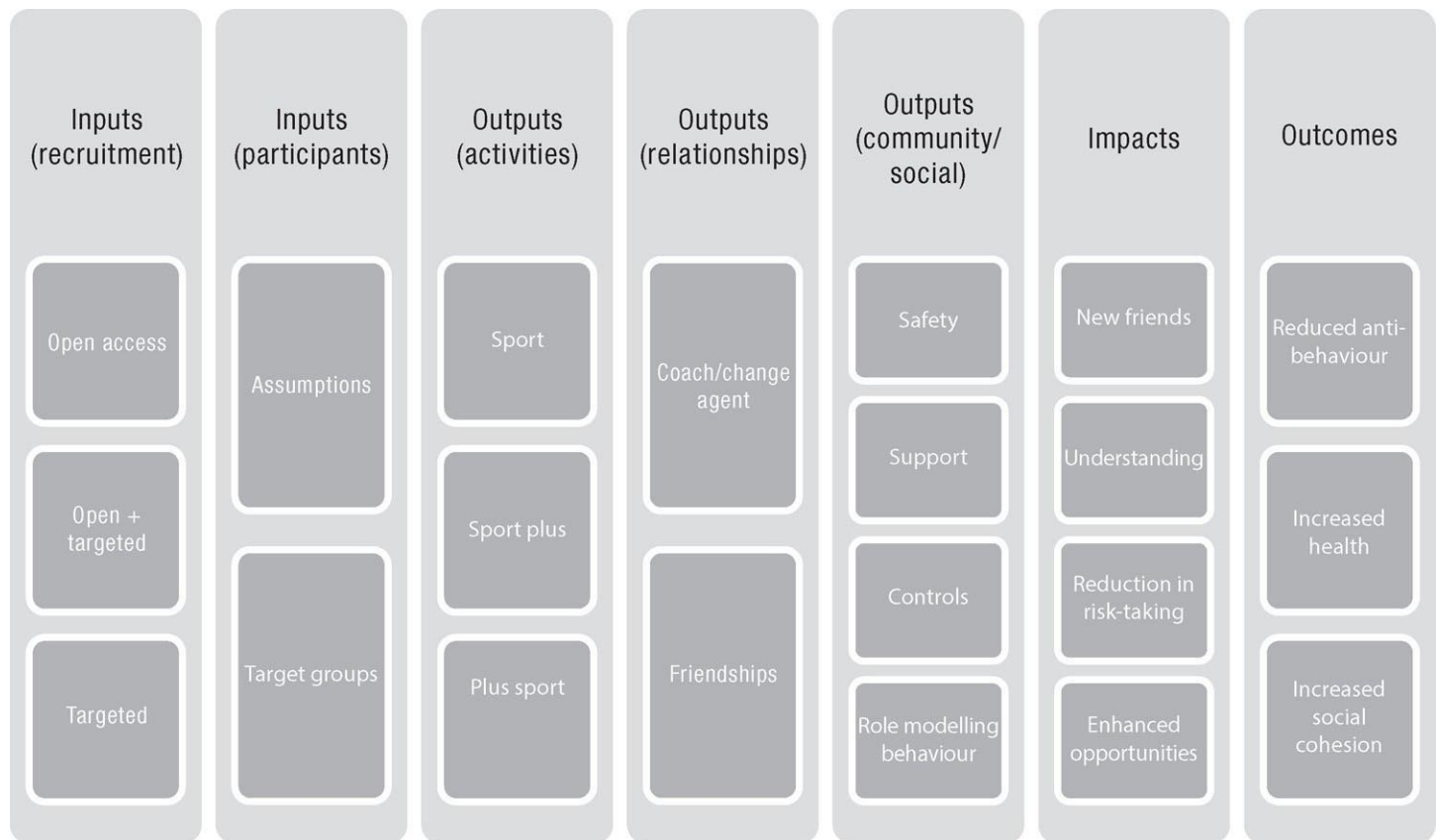
Figure 11.1 Program logic model

Source: Adapted from Coalter (2013a).

The first input is derived from the methods of *recruitment* (Coalter 2013a); these can range from open access, through self-selecting, to targeted and via referral. The program might be free, subsidised, or paid for by the participant. Each recruitment method will result in a different participant cohort; the recruitment method should be designed to garner the appropriate participants. Gathering sufficient information on the nature, extent, severity, and distribution of the issues that are the desired focus of the program determines the best recruitment method, which in turn ensures that the correct participants are recruited (Long 2011). In a SD context, the recruitment is more focused on specific talent identification or athlete pathways; however, the application of the program logic theory remains the same.

Further, gathering sufficient evidence on which to base program planning decisions is important to ensure that the program will be relevant, useful and have the desired outcomes for its intended participants (Houlihan 2011). There is a danger that organisations seeking to address an identified issue will impose their own understandings of a situation onto developing solutions (Coalter 2011). This is illustrated in the context of international SFD programs that are managed by staff from high income countries (HICs) seeking to address social, cultural, educational, or health-related issues associated with local communities in low- and middle-income countries (LMICs) (Schulenkorf 2012). In such cases, dominant ideologies – more relevant to the organisation's country of origin – may be imposed on program design and delivery. For example, Darnell (2011) demonstrated that interns working in Commonwealth Games Canada's International Development through Sport program understood their role as one that would facilitate the development of participants' leadership skills. The underpinning philosophy was that sport would facilitate personal development by creating individualised understandings of success and achievement for the participants, thereby leading to empowerment. Implicit in this philosophy is the notion that persons experiencing disadvantage have failed to achieve self-actualisation, an individualist notion that emanates from the neo-liberal philosophies of HICs and fails to account for the manifold pressures influencing an individual's development, such as war, poverty and disease (Darnell 2011). Researchers caution strongly against the wholesale imposition of neo-liberal philosophies, arguing that for sport for development programs to be effective – and to not just become another arm of global hegemony – design must begin in situ and with a proper assessment of the local understandings of desired outcomes (Coalter 2013a; Darnell 2011; Houlihan 2011; Schulenkorf *et al*. 2014).

The second input is *participants* (Coalter 2013a). Leading SFD scholars, such as Sugden (2010) and Coalter (2011, 2013a), warn against making assumptions about the inherent positive impact of sport and advocate the development of a clear

understanding of the context, nature, and distribution of the specific issues in focus. In the absence of adequate information and evidence about the issues that a program seeks to address, managers of programs run the risk of adopting a deficit view of participants that is based on environmental determinist assumptions that assume homogeneity in the cohort. Evidence- based planning ensures that measurable performance indicators are set around participants for both sport and non-sport outcomes.

Further, consideration of performance indicators facilitates decisions around what tools will be used to evaluate the program and its participants, as some outcomes are more suited to measurement by some tools than others (Houlihan 2011). For example, in an SD context, quantitatively measuring athlete physiological and skills metrics pre- and post-program may be a more effective method than qualitatively interviewing the athlete about how much body fat he or she thinks they might have lost or how much their vertical jump has increased. Similarly, in an SFD context, discovering the impacts of participation on an individual's development of social connections is better achieved through qualitative methods that allow discussion of what participation has meant to him or her in terms of their social situation. Tools for evaluation are discussed later in this chapter.

The first output varies depending on *the nature of the program* (Coalter 2013a), represented in Figure 11.1 by outputs (activities). For example, an SD program that offers sporting opportunities might be aimed at increasing participation in the sport, or developing mastery of the sport, or becoming proficient in the rules of the sport, or all three. On the other hand, an SFD program uses sport as the context, or the site, in which participants have the opportunity to gain desired skills or behaviours through their participation in the program. In both SD and SFD contexts, outcomes are derived not simply from the assumption that sport is good for you. Rather, it is the combination of the context and the specific activities included in the design of the program – such as attaining a coaching certificate or collaboratively developing a community sport event – that work towards achieving program outcomes. The choice of sport can be influential (Agans and Geldhof 2012); for example, an SFD program seeking to improve social skills might consider whether a team sport or an individual sport would be most effective for achieving the intended outcomes for the particular cohort.

The second output concerns *social relationships* (Coalter 2013a), represented in Figure 11.1 by outputs (relationships). The relationships with the personnel delivering the program, in combination with the type of program, have been shown to be a critical aspect of program effectiveness (Vella *et al*. 2013). In an SD context, the primary social relationship is with the coach; this relationship is typically constrained to sport-related activities and opportunities for developing closer personal relationships are limited. However, in an SFD context, the development of social relationships is often key to the success of the program (Coalter 2013b; Sherry 2010). For example, in the context of an SFD program seeking to address issues of gang membership, racism and at-risk youth, research has demonstrated that participants' outcomes are enhanced when trusting, respectful and reciprocal relationships are formed between participants and personnel (Coalter 2013b; Schulenkorf 2013).

The third output is the *social climate* (Coalter 2013a), represented in Figure 11.1 by outputs (community/social). This refers to the social environment in which the program is conducted. In an SD program, there might be an emphasis on understanding and abiding by the rules, learning cooperation or teamwork. In an SFD context, the social climate might be more directed towards creating a 'safe space' or inclusive environment in which the participants are comfortable and able to speak, listen and self-reflect (Spaaij and Schulenkorf 2014).

According to program theory, the impacts are about *providing a foundation for changing attitudes and behaviours* and *changing values and attitudes* (Coalter 2013a), the final column in Figure 11.1. In other words, the combination of the social relationships, the social climate and the activities create a context in which participants can develop the respect, trust and reciprocity required to achieve change. In an SD context, the foundation might be the combination of the acquired technical skills and adoption of the necessary attitude that enable the participant to perform at a high standard. In an SFD setting, this foundation of respect, trust and reciprocity might take the form of the participants seeking to not disappoint the coach, their team members or communities, and to conform to expectations. In terms of changing values and beliefs, varying with the nature of the program, the ability to achieve the desired changes depends on the inputs, outputs and on the receptiveness and suitability of the participants. SFD programs often seek to affect a change, for example, in attitudes towards people from other cultures. SFD programs thus emphasise the importance of understanding the implications of negative behaviours – both in a sporting context and elsewhere – in the hope that they will assume responsibility for their actions and adopt the desired changes.

Outcomes obviously vary depending on the nature of the program. An SD program might see the addition of a new team into a tournament or the development of a new league (see, for example, Swart *et al*. 2011); an SFD program might see a reduction in violence or antisocial behaviour (see, for example, Kelly 2013).

IMPLICATIONS OF THEORY FOR PRACTICE

Up to this point, this chapter has discussed SD and SFD programs as two interrelated yet distinct activities. This was done to demonstrate the variety in programming and the divergent outcomes intended from sport programming. However, it is helpful to think of SD and SFD as occupying either end of a continuum, as there are many similarities between the two. For example, some programs with an SFD focus may well serve to develop the sport, and programs seeking to develop the sport can also have positive community and personal development outcomes. Evaluating SD – that is, specific sport outcomes, can be focused on the individual athlete, and also broader community or national level outcomes. Common measurement indicators for SD are: athletes skill development and mastery, race and competition results, international rankings tables, and numbers of players and competitions. However, as noted above, SD can also argue, through appropriate research and evaluation, its contribution to public health outcomes such as increased cardio-vascular fitness, which can have a positive impact on good health across the community. Conversely, SFD evaluations tend to focus firstly on the non-sport outcomes being sought, such as education, work or inclusion, however with regular participation in sport activities, some sport development outcomes will also be achieved such as skill development and mastery. As a result, some participants in SFD programs may transition into SD programs, teams and competitions.

Monitoring and evaluation of a program's success is achieved largely through measurement of outcomes; however, as seen through an examination of program theory in this chapter, evaluation is threaded throughout the planning and design process. From identification of the program's objectives to recruitment of participants and program delivery, there is a constant need for MEL to determine the relevance and utility of program intent, delivery approach and desired outcomes. Rather than seeing MEL as something that occurs once the program is complete, evaluation needs to be embedded in design and considered throughout the process as it is influential in decision-making along the way. Evidence based decision-making is essential for both sport and non-sport outcomes, and equally applicable for SD and SFD programs and initiatives.

Evaluation, then, is both a *formative* part of the initial planning stages of programming and serving to shape and improve programming through amendment. It is also *summative*, providing useful feedback on completion of the program, to prove what was and was not successful (Houlihan 2011). Formative evaluation includes: *needs assessment*, in which the need for the program is assessed; *evaluability assessment*, to determine the feasibility of the evaluation; *structured conceptualisation*, in which stakeholders define the target population, the program, the approaches and methods for delivery, and possible outcomes; *implementation evaluation*, which monitors program delivery; and *process evaluation*, where delivery processes are monitored for possible improvements (Trochim 2006). Summative evaluation includes: *outcomes evaluation*, which investigates whether the program had demonstrable effects in the areas targeted; *impact evaluation*, which looks for broader effects of the program, intended and unintended; *cost-effectiveness and cost-benefit analysis* evaluates the efficiency of resource use; *secondary analysis*, in which alternative methods are considered; and *meta-analysis*, which brings together a range of research evaluating the same question (Trochim 2006).

Tools for evaluation

Consideration of evaluation practices and processes affects decisions about tools and research approach that will be taken to assess the outcomes of a program. As was mentioned earlier, there is a variety of research tools used in evaluation to determine the desired outcomes of a program which, in turn, assists in measuring the overall progress and success of a program. This section presents a brief overview of qualitative and quantitative tools used in evaluation.

Qualitative methods enable the voices of program participants to be heard (Corbin and Strauss 2008) and provide in-depth information about individuals' experiences (Thomas 2006). Qualitative tools include interviews, semi-structured interviews, and focus groups, as well as journaling and observation (Corbin and Strauss 2008). The focus of qualitative evaluation is on understanding, rather than generalizability, which is often the focus of quantitative evaluation (Creswell 2007). Qualitative methods often use purposive sampling, in which selection is made based on knowledge about the case or sample (Creswell 2007); this is to make sure that the garnered information is relevant and that the sample is sufficiently knowledgeable about the topic or issue under investigation.

In order to collect appropriate information, interview schedules – which outline what should be asked of participants – and research protocols – which provide structure regarding what information might be observed or collected – are developed with the gathering of specific information in mind (Kvale and Brinkmann 2009). In other words, interview schedules and research protocols should reflect the specific outcomes that are being measured. Vague or overly broad questions will likely encourage the sharing of information with limited utility or relevance to the issues under examination – hence, at least some structure to conducting interviews is encouraged (Kvale and Brinkmann 2009). Once data collection is

completed, qualitative evaluation typically uses inductive analysis in which detailed readings of raw data – for example, interview transcripts or researcher observation notes – are used to draw out themes, concepts, or a model (Thomas 2006). In this way, findings can be drawn from the data (Strauss and Corbin 1998), allowing evaluation of the programming from a participant's point of view.

Quantitative methods are well suited to the measurement of change. Causality and generalisability underpin quantitative evaluation, making it particularly useful for summative evaluation (Clarke and Dawson 1999). An SFD program aimed at reducing crime, for example, could measure the outcome in terms of crime statistics. On the other hand, an SD program might measure increases in participation rates or progress through development pathways. Quantitative techniques largely avoid research bias as the overall method design is facts-based and precludes the need for advanced interpretation (Punch 2014).

Quantitative tools include surveys, questionnaires, pre- and post-tests, and statistical information (Punch 2014); samples are typically much larger than those found in qualitative evaluation as the aim is to provide generalisable results (Long 2011). A range of quantitative instruments, scales and measures, designed to measure specific aspects of individuals' experiences, can be used in SD and SFD contexts. For example, the Positive Youth Development Inventory (Arnold *et al*. 2012) measures youths' psychological development and could be used pre- and post-program to evaluate outcomes of participation on participants' positive development.

Challenges in Research and Evaluation

Often MEL in sport – across both SD and SFD contexts – may be viewed either as a simple procedural exercise, or a rather easy administrative task. However, it is important to understand that MEL is at times more complex and requires thoughtful consideration in all settings. As such, when we are designing a program, we need to understand the answers to the following questions:

- What is the difference between MEL and research?
- What can we measure? And what can't we measure?
- Who conducts evaluations and who gets to see and/or use the findings?

First, it is also helpful to consider the difference between MEL and research. Monitoring can be best understood as a management tool to keep track of performance against set measures, such as a budget, attendance numbers or health and safety incidents. Evaluation is a management tool to understand why and how outcomes are being achieved, or not achieved, for example why attendance numbers may be dropping for a specific program, or which coaches are getting the best results with different training programs. Learning, in this context, refers to the knowledge gained and decisions made to subsequently refine or advance programs. Meanwhile, research generally presents a broader and less prescribed approach to investigating programs; it is generally undertaken by trained academics or researchers who are using specific and scientific approaches, drawing on sophisticated methodologies to understand complex questions such as behaviour change measures using psychometric surveys, or blood tests and muscle biopsies for high performance athletes.

Albert Einstein is quoted as saying "Not everything that can be counted counts and not everything that counts can be counted". Applied to our MEL and research context, quantitative assessments are generally those that can be counted – and hence properly measured – while qualitative data is less measurable in statistical terms. We need to be aware though that qualitative data is no less important or critical for the evaluation and advancements of programs – or, in Einstein's words, it still counts even if it cannot be counted.

The question of who gets to see or use the findings of any evaluation or research within the SFD and SD setting is often entwined with the socio-political context within which the programs are offered. For example, there is often a tension between the arms of SD and SFD regarding government funding, with investment decisions being made on athletic outcomes (e.g., medals at the Olympic Games) versus less tangible community outcomes (e.g., community physical and mental health from sport participation). If measurements of success are too narrowly focused, then things that "count" may not necessarily be "counted". The pressure to evaluate is also often embedded in funding agreements from governments, philanthropic donors, sponsors, and grant schemes. Although high quality evaluation can lead to effective evidence-based decision making, if the evaluation is required to simply meet funder requirements, or for the promotion of only positive outcomes, then it may lead to poor evaluation design or at worst provide inaccurate or misrepresentative findings. Evaluation and research are ideally undertaken by an independent body without undue influence on the approach or how the findings are to be presented or shared, to ensure that an honest and accurate assessment of the program or initiative can be used by all concerned to feedback into future program design and delivery with the organisation, its stakeholders and participants.

APPLICATION OF THEORY TO PRACTICE

The following case studies demonstrate evaluation in practice. The first case study presents an example of evaluation of a SFD program, while the second one focuses on a SD initiative.

Case Study 1: Timor Leste study

Katherine Raw

Timor Leste is a half-island nation with a population of 1.2 million, located north of Australia. Following 450 years of Portuguese occupation, Timor Leste was occupied by Indonesia for 24 years and, during that period, approximately one-quarter of the Timorese population were killed (Millo and Barnett 2004). In 2002, with the assistance of the UN, Timor Leste gained independence and is now considered to be one of the world's youngest countries (Millo and Barnett 2004; Robinson 2011). Since 2000, the UN, various NGOs and government groups have worked collaboratively to reduce poverty and violence, with the intention of promoting stability and progressive development throughout Timor Leste (McGregor 2007; Silove *et al.* 2009). However, with 53 per cent of the population under the age of 18 (UNICEF 2011), disengaged youth (Robinson 2011) and violence were key barriers to reducing social tensions and establishing peace (Goldsmith 2009; Scambary 2009).

In 2010, the Australian Catholic University (ACU) developed and implemented the Future in Youth (FIY) program in Timor Leste's second biggest town, Baucau. The program was coordinated in partnership with a handful of Baucau's local community leaders. Program development was centred on a community that had often been disrupted by hostility, primarily at the instigation of martial arts youth gangs, with high attrition rates from school, high levels of youth unemployment and disengaged youth. This SFD program used football as the sporting activity, and aimed to improve the health, well-being and life skills of youth and to build the capacity of the community. An additional long-term goal of the program is for the citizens of Baucau to have ownership of the program and sustain it independently of ACU. Assessing these particular objectives of the program was the primary focus of M&E activities that were supposed to be conducted regularly. However, by the time the FIY program was in its fourth year of operation in 2013, it had not instigated any formal M&E. The only form of assessment that had been undertaken was done directly by the program's Australian operators, largely in an informal manner. Therefore, in accordance with current recommendations from SFD research, the FIY program implemented M&E by engaging an independent researcher, not affiliated with the program. Community leaders and coaches involved with the FIY program (*n* = 24) were invited to share their views and opinions regarding the program and its impacts. Qualitative data were gathered through semi- structured interviews with community leaders and focus groups with coaches.

A question guide was developed to provide an overall framework summarising the content to be covered during discussions. As opposed to a more controlled approach, the question guide was developed with the idea of allowing questions to be asked in any order with minimal fixed wording and answered with various time allowances (as per Morgan and Guevara 2008). To reduce the potential for prompting or bias, the FIY program's aims were not directly included within the question guide; however, if participants broached these topics, they were allowed to discuss them. The final question guide for this study comprised several open-ended questions asking participants about their opinions regarding the FIY program, with the option of using more targeted questions to explore topics as necessary. Once finalised, the primary interpreter translated the question guide into the local language of Tetum; it was then two-way translated from Tetum back into English – without prior viewing of the English questions – by a second interpreter in order to minimise the potential for concepts to become lost in translation (Muthoni and Miller 2010; Squires 2009).

While not specifically a part of formal data collection, reflexive journaling was also undertaken to enhance the depth of data that emerged through semi-structured interviews and focus groups. These processes occurred at the completion of the FIY program, and again three months later to examine whether the program was achieving and sustaining its identified outcomes. Following the completion of interviews and focus groups, data were translated, transcribed and pooled. Analysis of data was done using thematic analysis techniques, which involved identifying, analysing and then reporting on themes that emerged from the data. A hierarchal structure of themes emerged, allowing the discovery and investigation of patterns in the data, which were related back to the question guide and therefore, in turn, assessed in relation to the research questions. To enhance the dependability of data analysis, two researchers undertook this process separately and compared the emergent themes.

Positive links were found between the FIY program and the health, well-being and life skills of the program's youth participants, coaches and community leaders. To an extent, sport-specific community capacity was also developed through football and coaching skills. The findings regarding program sustainability were mixed, with the majority of participants expressing that the program was sustained for four years in a largely unstructured manner by the youth who had participated in the program with little input from the coaches. However, there were some community leaders and coaches who indicated their willingness to take more responsibility to ensure the program would be sustained in the future. This particular point was instrumental in moving the FIY program forward in terms of sustainability, as these individuals have since become key facilitators of improvements in local program sustainability in 2014.

The results gathered through the initial bout of M&E in 2013 highlighted the value and necessity of the evaluation process to program organisers and funders. This understanding was solidified again in 2014 with a second year of M&E that produced an even greater depth of information and honesty from participants. As a result, ongoing M&E was implemented the following year and will continue to be a permanent practice into the future.

Further resources
Evaluation and Program Planning (an open-access Elsevier journal devoted to evaluation research):
www.sciencedirect.com/science/journal/01497189
Innovation Network 2014, Point K: Practical Tools for Planning, Evaluation and Action:
www.innonet.org/index.php?section_id=4&content_id=16

Case study extension activities
1. What are the possible benefits and disadvantages of having M&E undertaken by independent researchers?
2. In what ways do you think evaluation research outcomes were served by the data being examined by two independent researchers?
3. Why do you think the M&E process was even more enlightening in the second year it was conducted?
4. How might you incorporate opportunities for M&E into program design?

Case Study 2: Hockey Victoria
Angela Osborne

Hockey Victoria (HV) is a state body charged with oversight of hockey in the Australian state of Victoria. Hockey is a mixed-gender sport with some 22,000+ members registered in Victoria (HV 2014). HV had noted that the participation rate for girls was steadily dropping and that girls in the 16–19-year-old bracket were quitting the sport, resulting in a talent gap beginning to form in the elite levels as insufficient numbers of girls were going on to play senior hockey. In 2010, HV received funding from state government body VicHealth to put into place a suite of inclusion programs aimed at improving the participation and retention rates of women and girls in hockey (HV 2014).

In addition to improving retention and increasing participation rates of junior girls, the program sought to raise awareness of the contributions of women and girls in hockey, both in the sport and in individual clubs. To achieve these strategic aims, HV developed a range of programs, each of which was directed towards addressing specific issues, with specific outcomes in mind.

Inclusion awareness training was offered to HV staff and board and was intended to raise awareness of the issues facing diverse communities within the sport of hockey and to garner broad organisational support for the inclusion strategy. The pilot club program brought several clubs together regularly to discuss issues and solutions; a range of speakers offered club representatives ideas and information on how to improve the participation of girls and women. Clubs developed different kinds of initiatives and discussed their implementation; HV provided support and resources for clubs seeking to affect change in their club. In addition, HV developed and delivered a range of courses and programs aimed at addressing specific aspects of the identified issues. A coaching program was established that offered girls and women an opportunity to learn new coaching skills and to get their coaching certificate. This program was developed with the intent of raising the visibility of women and girls in hockey. HV found that girls can be intimidated by male officials, which made them unlikely to seek assistance for problems and had an indirect influence on decisions to leave the sport. The coaching program was a long-term strategy to achieve gender balance in officials with a view to improving retention.

A junior leadership program was developed with the intent of establishing and consolidating girls' sense of ownership and belonging within their club. This program sought to empower girls to encourage them to remain in the club through the

critical 16– 19-year-old period. A governance workshop offered club representatives ways to address inequity in their club at the governance level; for example, by ensuring that the board was representative of club membership – it had been noted that many clubs had very few, or no, female board members. A coaching workshop devoted to coaching girls was offered. Coaches and club representatives were provided with training about gender differences and on how to adapt their coaching practices to ensure that girls feel, and are, included as valuable members of the team.

A mentoring program was established that matched current junior elite players with older, more experienced elite players. This program sought to provide elite athletes with an avenue for discussion about their hockey and life experiences. The underpinning intent was to offer additional support to elite athletes – which would not impact on selection or standing within their clubs – with the intent of encouraging retention and transition into senior elite hockey. Evaluation of the suite of programs was undertaken halfway through the three-year funding period and again at the end of the funding period. Mixed methods were used to evaluate different aspects of the suite of programs. Qualitative evaluation questioned program participants' understandings of inclusion, their experiences in the specific program they were involved with, how their clubs had changed following participation and recommendations for further improvement. Overall, the qualitative evaluation sought to ascertain the extent to which the principles underpinning the programming had been understood and embedded into practice; thus, qualitative evaluation was necessary to properly understand the range and depth of the impact on the hockey community.

For the qualitative evaluation, an interview schedule was developed that addressed the research questions and encouraged discussion of the aspects of the programs related to increasing visibility and improving retention of women and girls in hockey. The questions on the interview schedule were all crafted to garner information about a specific aspect of the programming and its underpinnings. For instance, in order to assess the adoption of the principles of inclusion, research participants were asked to discuss their opinions on creating a safe, welcoming and inclusive environment. Responses to this question provided information about the extent to which the ideas underpinning the programs had been understood and incorporated into practice; responses offered information about possible gaps in understanding, which can then be used to develop further programs. Participants for the qualitative evaluation included program participants, HV staff and the HV board. HV provided a complete list of program participants, all of whom were invited to take part in the evaluation. Contact was initially made via email; the response rate to the invitation was 73 per cent. Research participants provided their phone contact details and a researcher conducted a telephone interview at a time of their choosing. Interviews were recorded and transcribed. A qualitative software management program, NVivo, was used to manage and sort the data; thematic analysis was used to uncover themes within the data.

Through the process of qualitative evaluation, HV received information about the impact and outcomes of programs, with which it was then possible to determine the success of the programming. For example, evaluation revealed that: the pilot club model was successful and had begun to have results, with an increase in the numbers of junior girls noted across pilot clubs; women and girls felt that their status within clubs had improved; and the focus of the suite of leadership, governance and coaching programs was having a positive impact on retention as girls felt supported, encouraged and empowered to continue their hockey participation through the critical period. Moreover, evaluation undertaken midway through the three-year funding period enabled HV to adjust programs to better meet the needs of participants. For example, evaluation of the mentor program midway revealed that participants who were geographically distant struggled to meet, thereby reducing the effectiveness of the program. Amendments were made on the strength of this finding, and subsequent participants were paired with geographic proximity in mind. Similarly, midway evaluation found that some mentors – particularly those employed in a coaching role – felt unsure about the precise nature of the relationship; this information allowed HV to provide clarification about the nature of the role and how it differed from that of a coach. The second cohort of mentors was clearer about their role and no such problems were encountered.

The summative evaluation provided information on how to better HV programming. For instance, evaluation of the governance workshop revealed that while regional clubs were benefited by the inward-looking, club-focused governance activities, metropolitan clubs would have preferred an outward-looking, whole-of-sport focus that examined strategic direction for the sport. Metropolitan clubs, by virtue of their location and surrounding populations, had been concerned with inclusion – either directly or indirectly – for some time, meaning that their needs around governance and inclusion differed markedly from those of clubs in regional areas. This revelation enabled HV to develop programming more suited to the needs of the participants. Further, the finding provided HV with crucial information delineating the experiences of metropolitan and regional clubs, enabling better understanding of how programming and strategy might be developed in a more targeted fashion that accounts for these variations in experience in the future.

Further resources

Hockey Victoria (HV): www.hockeyvictoria.org.au/

Case study extension activities

1. Why do you think the evaluations adopted a qualitative approach?
2. Why would it be important for HV to garner an in-depth understanding of program participants' experiences?
3. How might quantitative approaches have been used?
4. What would be the benefits and challenges of each kind of evaluative approach for M&E of HV's suite of programs?

SUMMARY

This chapter has examined the role of MEL and research in sport programming. It has demonstrated that MEL processes are necessary and useful at all stages of planning, implementation, and delivery in order to ensure program outcomes are relevant, useful, and successful for intended cohorts. The chapter has provided a framework with which to embed evaluative practice in program design and has demonstrated the commonalities and differences between MEL in SFD and SD settings. It has also highlighted the challenges and considerations inherent in MEL and research across the two arms of sport development.

DISCUSSION QUESTIONS

1. Identify and discuss the difference between quantitative, qualitative and mixed- method research approaches.
2. You are developing an SD program aimed at developing individual athletes in a specific sport. What should be considered in the research design process when developing an evaluation plan? List the factors and state why these are important.
3. Imagine you are developing an SFD program aimed at reducing truancy in your city or region. What should be considered in the research design process when developing an evaluation plan? List the factors and state why these are important.
4. Why do funders and stakeholders require monitoring and evaluation to be undertaken for sport programs? What are the possible implications of failure to do so?
5. Evaluation can be undertaken within the program or by external experts – discuss the pros and cons of each approach.

REFERENCES

Agans, J. P. and Geldhof, G. J. (2012) 'Trajectories of participation in athletics and positive youth development: the influence of sport type', *Applied Developmental Science*, 16(3): 151–165.

Arnold, M. E., Nott, B. D. and Meinhold, J. L. (2012) *The Positive Youth Development Inventory (PYDI)*, Corvallis, OR: Oregon State University 4-H Youth Development Program.

Clarke, A. and Dawson, R. (1999) *Evaluation Research: An Introduction to Principles, Methods and Practice*, London: Sage.

Coalter, F. B. Houlihan and M. Green (2011) 'Sport development's contribution to social policy objectives: the difficult relationship between politics and evidence', in B. Houlihan and M. Green (eds), *The Routledge Handbook of Sports Development*, London: Routledge, pp. 561–578.

Coalter, F. (2013a) *Sport for Development: What Game Are We Playing?* London: Routledge.

Coalter, F. (2013b) '"There is loads of relationships here": developing a programme theory for sport-for-change programmes', *International Review for the Sociology of Sport*, 48(5): 594–612.

Corbin, J. and Strauss, A. (2008) *Basics of Qualitative Research*, Thousand Oaks, CA: Sage. reswell, J. W. (2007) *Qualitative Inquiry and Research Design: Choosing among Five Approaches*, Thousand Oaks, CA: Sage.

Darnell, S. C. (2011) 'Identity and learning in international volunteerism_ "sport for development and peace" internships', *Development in Practice*, 21(7): 974–986.

Goldsmith, A. (2009) '"It wasn't like normal policing": voices of Australian police peacekeepers in Operation Serene, Timor-Leste 2006', *Policing and Society*, 19(2): 119– 133.

Hockey Victoria (HV) (2014) *Annual Report*, available at:
www.hockeyvictoria.org.au/Portals/15/2014%20Media%20Uploads/2014HVAnnualRep Website.pdf (accessed July 2015).

Houlihan, B. B. Houlihan and M. Green (eds), B. Houlihan and M. Green (2011) 'Introduction: the problems of policy evaluation', in B. Houlihan and M. Green (eds), B. Houlihan and M. Green (eds), *The Routledge Handbook of Sports Development*, London: Routledge, pp. 557–560.

Kelly, L. (2013) 'Sports-based interventions and the local governance of youth crime and antisocial behavior', *Journal of Sport and Social Issues*, 37(3): 261–283.

Kvale, S. and Brinkmann, S. (2009) *InterViews: Learning the Craft of Qualitative Research Interviewing*, 3rd edn, Los Angeles, CA: Sage.

Long, J. K. Hylton and P. Branham (2011) 'Researching and evaluating sports development', in K. Hylton and P. Branham (eds), *Sports Development: Policy, Process and Practice*, 2nd edn, London: Routledge, pp. 236–257.

McGregor, A. (2007) 'Development, foreign aid and post-development in Timor-Leste', *Third World Quarterly*, 28(1): 155–170.

Millo, Y. and Barnett, J. (2004) 'Educational development in East Timor', *International Journal of Educational Development*, 24(6): 721–737.

Morgan, D. and Guevara, H. L. Given (2008) 'Interview guide', in L. Given (ed.), *The Sage Encyclopedia of Qualitative Research Methods*, Thousand Oaks, CA: Sage, pp. 470– 471.

Muthoni, H. and Miller, A. (2010) 'An exploration of rural and urban Kenyan women's knowledge and attitudes regarding breast cancer and breast cancer early detection measures', *Health Care for Women International*, 31(9): 801–816.

Punch, K. F. (2014) *Introduction to Social Research: Quantitative and Qualitative Approaches*, 3rd edn, London: Sage.

Robinson, G. (2011) 'East Timor ten years on: legacies of violence', *The Journal of Asian Studies*, 70(4): 1007–1021.

Scambary, J. (2009) 'Anatomy of a conflict: the 2006–2007 communal violence in East Timor', *Conflict, Security & Development*, 9(2): 265–288.

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.

Schulenkorf, N. (2013) 'Sport-for-development events and social capital building: a critical analysis of experiences from Sri Lanka', *Journal of Sport for Development*, 1(1): 25–36.

Schulenkorf, N., Sugden, J. and Burdsey, D. (2014) 'Sport for development and peace as contested terrain: place, community, ownership', *International Journal of Sport Policy*, 6(3): 371–378.

Sherry, E. (2010) '(Re)engaging marginalized groups through sport: the Homeless World Cup', *International Review for the Sociology of Sport*, 45(1): 59–71.

Silove, D., Brooks, R., Bateman Steel, C. R., Steel, Z., Hewage, K., Rodger, J. and Soosay, I. (2009) 'Explosive anger as a response to human rights violations in post-conflict Timor-Leste', *Social Science & Medicine*, 69(5): 670–677.

Spaaij, R. and Schulenkorf, N. (2014) 'Cultivating safe space: lessons for sport-for- development projects and events', *Journal of Sport Management*, 28(6): 633–645.

quires, A. (2009) 'Methodological challenges in cross-language qualitative research: a research review', *International Journal of Nursing Studies*, 46(2): 277–287.

Strauss, A. L. and Corbin, J. M. (1998) *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, Thousand Oaks, CA: Sage.

Sugden, J. (2010) 'Critical left-realism and sport interventions in divided societies', *International Review for the Sociology of Sport*, 45(3): 258–272.

Swart, K., Bob, U., Knott, B. and Salie, M. (2011) 'A sport and sociocultural legacy beyond 2010: a case study of the Football Foundation of South Africa', *Development Southern Africa*, 28(3): 415–428.

Thomas, D. R. (2006) 'A general inductive approach for analyzing qualitative evaluation data', *American Journal of Evaluation*, 27(2): 237–246.

Trochim, W. (2006) *Research Methods Knowledge Base*, available at: www.socialresearchmethods.net/kb/index.php (accessed July 2015).

NICEF (2011) 'At a glance: Timor-Leste', *UNICEF*, available at: www.unicef.org/infobycountry/Timorleste_statistics.html (accessed July 2015).

Vella, S. A., Oades, L. G. and Crowe, T. P. (2013) 'The relationship between coach leadership, the coach-athlete relationship, team success, and the positive developmental experiences of adolescent soccer players', *Physical Education and Sport Pedagogy*, 18(5): 549–561.