

Commentary

Sport and Social Inclusion: Evidence-Based Policy and Practice

Fred Coalter^{1,2}

¹ Carnegie School of Sport, Leeds Beckett University, Leeds, LS1 3HE, UK; E-Mail: fredcoalter@hotmail.co.uk

² Department of Movement and Sport Sciences, Vrije Universiteit Brussel, 1050 Brussels, Belgium

Submitted: 2 January 2017 | Accepted: 25 January 2017 | Published: 29 June 2017

Abstract

This commentary reflects on my experience of compiling the *Value of Sport Monitor*—an on-line resource of policy-relevant, research on the social impacts of sport—for eight years. The commentary critically evaluates the assumption of the *Value of Sport Monitor* that social science research in sport is cumulative and it explores sports interest groups' varying attitudes to the nature of evidence. It illustrates that widespread conceptual and methodological inconsistencies and weaknesses in research greatly reduce the ability to identify best practice and 'best buys' as a basis for policy. The commentary concludes by proposing that a way forward for research to contribute to policy and practice is via theory-based evaluation.

Keywords

evidence-based policy; methodological weaknesses, sports research; theory-based evaluation

Issue

This commentary is part of the issue "Sport for Social Inclusion: Questioning Policy, Practice and Research", edited by Reinhard Haudenhuyse (Vrije Universiteit Brussel, Belgium).

© 2017 by the author; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

This commentary is based on my experience of compiling the *Value of Sport Monitor* between 2004 and 2012. This was funded by Sport England and UK Sport and was an on-line data base of English language policy-relevant sports research covering the following topics (viewed as covering most aspects of the amorphous notion of 'social inclusion'):

- Physical Fitness and Health
- Psychological Health and Well-being
- Crime Reduction and Community Safety
- Economic Impact and Regeneration of Local Communities
- Education and Life-long Learning
- Social Capacity and Cohesion
- Influences on Participation

A basic assumption of the monitor was that research in sport was cumulative. The task was to identify research which proved the claims that sport interest groups make about its contribution to the solution of public problems

and aspects of social inclusion. It also aimed to identify best practice or 'best buys' as guides for policy and practice—in terms of the 'New' Labour mantra, to identify 'what works'. It was also hoped that it would identify gaps in the research, thereby encouraging researchers to undertake work to fill the gaps. Consequently, the monitor was not a bibliography but consisted of narrative reviews and regular state-of-the-art summaries, identifying gaps in research.

The context for the establishment of the monitor was the 'New' Labour government's twin emphasis on social inclusion and evidence-based policy making. 'New' Labour's Third Way concerns with social inclusion, civic renewal, active citizenship and social capital were reinforced by an emphasis on joined-up government and a policy imperative that all areas of public investment contribute to strategic policy goals. This placed sport on a broader policy agenda because it was impossible for government departments, non-departmental public bodies (e.g., Sports Councils) and those in receipt of public funding to ignore it. Funding became increasingly dependent on an organisation's ability to illustrate its effective contribution to the social inclusion agenda.

For example, the 1999 white paper *Modernising Government* (Cabinet Office, 1999) stated ‘this Government expects more of policy makers...better use of evidence and research in policy making and better focus on policies that will deliver long term goals’.

In such circumstances the relatively untested claims of sport’s wider social contributions came under much closer scrutiny and robust research to inform evidence-based policy-making was in short supply. For example, in 2002 the ‘New’ Labour government’s sports strategy *Game Plan* (Department of Culture, Media and Sport & Strategy Unit, 2002, p. 79) stated that ‘the evidence base needs to be strengthened to enable policy-makers to construct and target effective interventions’. In a review of evidence of the socio-economic benefits of sport participation for the Conference Board of Canada, Bloom, Grant and Watt (2005, preface) concluded that ‘policy makers lack the evidence required to make informed policy decisions and to connect sport issues to other priorities’.

Such comments illustrate that another rationale for the Monitor was to add legitimacy to what was a relatively marginal policy area. For example, Houlihan and White (2002) argue that sports development has mostly been subordinate to much more powerful policy communities (e.g., health and education) and it has been a policy taker not policy maker. Consequently, it could be argued that the social inclusion agenda was a mixed blessing, simply reinforcing this status—a shift away from the social democratic welfare vision of ‘sport for all’ to a much more pragmatic emphasis on ‘sport for good’ and accountability. This was starkly illustrated by the Minister, Richard Caborn’s (2003) assertion that ‘we will not accept simplistic assertions that sport is good as sufficient reason to back sport’. The increased emphasis on evidence of effectiveness can also be viewed as part of the struggle to establish legitimacy in the eyes of other established policy fields (e.g., health; crime), dominated by high status professionals, with an assumed accumulated body of systematic and ‘scientific’ knowledge, who express scepticism about sport’s claims to funding related to their policy areas.

2. Interest Groups and Attitudes to Evidence

Nick Rowe (2005), the Research Officer at the Sports Council, identified three groups with an interest in the monitor. These were government, sporting organisations and academic researchers, each with different attitudes to the nature of evidence and proof.

2.1. Government: Truth and Utility

For Rowe (2005), government was looking for short-term definitive answers relating to the economy and effectiveness of policy-interventions. Government is characterised by polarised views and ‘evidence’ is evaluated as to the extent to which it supports and reinforces policy beliefs and current policy commitments. Weiss and Bu-

cuvalas (1980), in a series of 52 in-depth interviews with senior government officials to explore the usefulness of mental health research, found that decision-makers used a series of frames of reference to assess both the ‘truth’ and the ‘utility’ of social science research. Although the quality of the research was a major concern, Weiss and Bucuvalas (1980, p. 304) found that, ‘even more pervasive is the sense of conformity with what they already know—or believe they know’. Consequently, findings which reinforce beliefs and policies and are congruent with habits and tradition are more likely to be viewed as ‘useful’. This reflects Weiss’ (1993, p. 94) contention that evaluation is ‘a rational exercise that takes place in a political context’. In this world ‘evidence’ is more plural than research—values and ideology, habits and traditions, lobbyists, experience, expertise and judgement of officials and ministers all play their part in mediating and interpreting the value of research evidence. Weiss (1993, p. 96) emphasises the importance of political judgements in assessing research evidence by arguing that ‘a considerable amount of ineffectiveness may be tolerated if a program fits well with prevailing values, if it satisfies voters, or if it pays off political debts’.

In this regard Solesbury (2001) warns against the danger of academic researchers confusing research with evidence and believing that only academic research counts in policy and decision-making. Solesbury (2001, p. 9) argues that:

Public policy is developed and delivered through the use of power....There sometimes seems to be a tension between power and knowledge in the shaping of policy....Emphasising the role of power and authority at the expense of knowledge and expertise in public affairs seems cynical; emphasising the latter at the expense of the former seems naïve.

For these reasons, Weiss and Bucuvalas (1980) argue that there are very few ‘big bang’ moments in social policy, when a piece of research—a killer fact—overturns years of accumulated policies and practices. Because of the political, bureaucratic and interest-ridden nature of organisational decision-making processes, there are few examples of direct and immediate influence of research on decisions. Consequently, the main purpose of the monitor for government was, where possible, to confirm policy commitments already made, or to provide civil servants and ministers with positive examples.

2.2. Sports Agencies: Making the Case

The second set of interest groups is sports agencies and organisations—whose aim is to promote public investment in sport. These are characterised as being lobbyists and enthusiasts with an a priori belief in the benefits of sport. This group of sports evangelists is interested in evidence to ‘make the case’ and has limited interest in research quality/ methodology. In such circumstances

evidence is often selective and less systematic—e.g., using single research studies to argue a case—with limited concern about the robustness and limits to generalisability of the chosen examples. The danger in ignoring issues of method is emphasised by Oakley, Gough, Oliver and Thomas (2005) who found that in general social science research, better designed studies were generally less likely to demonstrate effectiveness, with poorly designed studies producing over-optimistic results. In the area of sport, Etnier et al. (1997), reviewing the evidence relating to the relationship between physical activity and educational performance, found that the largest measured relationships are obtained from the weakest research designs and the weakest relationships are found in the most robust research designs.

It could be argued that the approach to evidence in sport reflects its 'mythopoeic' nature. Mythopoeic concepts tend to be ones whose demarcation criteria are not specific and this can be applied to the way that an overly-homogenised notion of 'sport' is used in policy debates and rhetoric. Such concepts are based on popular and idealistic ideas which are produced largely outside sociological analysis and which 'isolates a particular relationship between variables to the exclusion of others and without a sound basis for doing so' (Glasner, 1977, pp. 2–3). Such myths contain elements of truth—some sports programmes will inevitably produce positive outcomes. However, these elements become reified and distorted and 'represent' rather than reflect reality, standing for supposed, but largely unexamined, impacts and processes. The strength of such myths lies in their 'ability to evoke vague and generalised images' (Glasner, 1977, p. 1)—useful in the rhetoric of 'sport' policy and lobbying. As Long and Sanderson (2001) argue, such a belief in the positive possibilities of sport is maintained in the absence of robust confirming evidence—a mixture of belief and theory, professional and personal repertoires, political and organisational self-interest and ad hominem arguments permits the assumption of such outcomes.

This combination of evangelical belief in the mythopoeic power of sport and a desire for legitimating evidence raised significant issues for the role of negative findings in the monitor. An implicit, but mistaken, assumption underpinning the notion of cumulative evidence is that the research will largely illustrate positive outcomes and impacts. Given the diversity of contexts, programmes and participants it was inevitable that not all research reported positive outcomes and in some areas, such as crime, there were equal proportions of positive and negative findings. However, learning from failure poses dilemmas for a vulnerable policy area, which is under pressure to deliver short term outcomes in a relatively short window of opportunity. In such circumstances, public agencies may be reluctant to publish negative results. The potential significance of this is emphasised by Sibley and Etnier (2003, p. 253) who, in examining the relationship between physical activity and cognition in children, conclude that, 'studies

with null results are often not published...which typically leads to a positively biased effect size being found for published studies'. Consequently, the inclusion of negative findings was the subject of much debate and such findings were included only where they contributed to our overall understanding of the issues, e.g., by explaining the causes of the failure of programmes to achieve their desired outcomes.

2.3. Academics

Rowe's policy maker's ideal type of this group is that academics are interested in the robustness of research design and methodology and concerned with issues of causality and theoretical explanations—concerns reinforced by the peer review process in journals. He also contends that academics are more likely to adopt a neutral 'enlightenment' approach, being more concerned with a critique of current policy and practice rather than a partisan, problem-solving approach. While this may be an over-generalisation, it is a position widely held in politics and by policy makers. For example, David Blunkett (2000, p. 15), a government minister, argued that:

Too much social science research is inward looking, too piecemeal rather than helping to build knowledge in a cumulative way....Issues for research are too supplier driven rather than focusing on the key issues of concern to policy makers, practitioners and the public at large.

Perhaps such perspectives explain the increasing employment of consultants in many areas of policy research (including sport), as they are perceived to be more pragmatic and user-friendly than academic researchers, who are often viewed as inflexible and naïve about client needs.

3. Weaknesses of Current Research

The needs of evidence-based policy making were not well-served because of the lack of a strong cumulative body of social science research evidence from which to inform sport policy and practice. This was explained by three broad factors: conceptual variety, methodological weaknesses and failure to address issues of sufficient conditions. It is important to note that, although this commentary concentrates on sports research, the weaknesses and limitations which will be highlighted are not confined to sport—evidence-based policy making has exposed widespread limitations in many areas of research, especially its ability to inform policy (Faulkner, Taylor, Ferrence, Munro, & Selby, 2006; Oakley et al., 2005). Davies (2004, p. 21) has stated that various research syntheses 'have shown that there is often a lack of sound conclusive evidence even when there has been considerable research activity on some topic or problem'. Consequently, many of the issues raised in this commentary

relate to generic issues of social science research and its contribution to policy.

3.1. Conceptual Variety

Weiss (1993, p. 96) has argued that marginal policy fields seeking to establish their legitimacy may make ‘inflated promises [with] goals lacking the clarity and intellectual coherence that evaluation criteria should have’. In this regard, Harris and Adams (2016, p. 2) argue that, ‘many problems with which sport is charged with “fixing” are poorly defined, lack clarity and are resistant to clear and agreed solutions’. This lack of clarity is reflected in much of current social science research. Not surprisingly, there was more cumulative evidence in terms of outcome measurement in medical research, on physical and psychological health, where there is more conceptual consensus and agreement on approaches to measurement. Much available social science research exhibits a wide variety of frequently vague definitions of ‘sport’ and this often included physical activity. These varying definitions often ignored the wide variety of processes and participants’ experiences—e.g., rule-bound organised sport is different from the more general category of ‘physical activity’. For example, in a study of sport and character development the President’s Council on Physical Fitness and Sports (2006) argued that it was not possible to generalise about sport because of the variety of rule structures, developmental stimuli, sporting cultures and micro-cultures and moral norms and the variety of individual experiences.

Jones, Edwards, Bocarro, Bunds and Smith (2017) in an integrative review of USA sport-based youth development literature found that only 55 percent of articles reported the specific sports being studied, only 10 percent reported the classification of sports being studied (i.e., team vs individual) and 35 percent treated sports participation as a single variable (i.e., participation /non-participation); much published research lacks detail of the nature, extent and duration of participation required to achieve certain outcomes—in Jones et al.’s (2017) study only 28 percent reported duration and 17 percent reported frequency; there was a wide variety of often disputed definitions of attitudinal and psychological outcomes and impacts (the resulting individual behavioural changes). For example, the varied measurements of educational performance greatly reduced validity and comparability— self-report, teacher evaluation, self-assessment and objective measurement. In much-researched areas there were vague and imprecise notions of ‘anti-social behaviour’, ‘at risk’ youth and ‘social cohesion’ and this raised issues of validity and reduced comparability and cumulative understanding. Much of the work on social cohesion was at a developmental stage and the definition and measurement of sport’s contribution to aspects of social cohesion and social capital present researchers with considerable methodological difficulties. The research in this area addressed is-

ues at various material and conceptual levels: large scale sports development programmes; the community impact of amateur and professional sports clubs; the role of sport and volunteering in developing pro-social behaviours and organisational cohesion; sport’s role in developing peer relationships and the meaning of sport for minority ethnic groups with differing attitudes to social and cultural integration. In terms of economic impact the variation in definitions and the variable data sources placed substantial limitations on comparisons in many areas—especially health-related savings and macro-economic impacts.

Such variety and lack of precision raise substantial issues of validity and comparability of research findings and substantially reduces the possibility of cumulative research understanding.

3.2. Methodological Issues

One effect of the enlightenment orientation is that the focus of much published research is the measurement of outcomes and the development of the methodology and the discipline—reinforced by the requirements of many journals and the peer review process. Such papers are of academic interest but it is very difficult to assess the practical relevance or lay accessibility of much of this type of research. The inclusion of necessary methodological caveats often undermines attempts to summarise the research for policy makers, as caveats can often outweigh evidence. However, many of the caveats are related to generic issues of method and could have been foreseen before the research was undertaken—cross-sectional designs; convenience sampling; self-selecting participants; no control groups; self-reporting; lack of control for intervening/confounding variables. A substantial number of academic articles finish with methodological caveats about the inherent limitations of the chosen methodology and such issues raise substantial questions of reliability. Further, there has been a dramatic increase in the volume of such research, reflecting the expansion of the publish-or-perish culture in universities. The pressure to publish has led to ‘salami-slicing’ and the failure to publish all relevant detail in one place. All this makes communication with policy makers and practitioners very difficult.

For example, in a review of correlates of physical activity of children and adolescents Sallis, Prochaska and Taylor (2000) looked at 108 studies which identified 40 variables for those aged 3–12 and 48 variables for the 13–18 age group—60 percent of the associations were statistically significant! The most notable result was a lack of consistency between the variables, with few variables consistent in all studies. The key constraints on the reviewed studies included:

- Low sample sizes
- Difficulties in measuring young people’s physical activity
- Differing sample characteristics

- Different analysis strategies: bi- and multivariate
- Cross-sectional studies

More generally, it is interesting to note how limited our understanding is of a most basic question—why do people participate in sport? Or, how do we achieve ‘sporting inclusion’? At the most basic level of understanding of the determinants of participation there is a lack of robust research evidence to inform policy and provision. For example, Biddle, O’Connell and Braithwaite (2011) in a review of quantitative systematic reviews concluded that ‘beyond age and gender, correlates are likely to have only small or small-to-moderate effects in isolation and may work best in interaction with other influences, although we are not close to identifying the *nature of these interactions*’.

Jackson, Howes, Gupta, Doyle and Waters (2005) undertook a systematic review entitled “Interventions Implemented through Sporting Organisations for Increasing Participation in Sport”, using restrictive selection criteria, including randomised and cluster controlled trials, quasi-randomised trials and controlled before-and-after studies. They could find no relevant studies and concluded that:

There is an absence of high quality evidence to support interventions designed and delivered by sporting organisations to increase participation in sport. Interventions funded and conducted in this area must be linked to a rigorous evaluation strategy in order to examine overall effectiveness, socio-demographic differentials in participation and cost-effectiveness of these strategies. (Jackson et al., 2005, p. 2)

Much is made of sports’ ability to contribute to the social and emotional well-being of ‘at-risk’ youth (although ‘at risk’ is rarely defined precisely). However, in a systematic review Lubans, Plotnikoff and Lubans (2012, p. 2) concluded that ‘the quality of existing studies is poor and has not improved since earlier reviews’. They state: ‘Due to the mixed findings and the high risk of bias, it is difficult to determine the efficacy of physical activity programmes for improving social and emotional well-being in at-risk youth’ (Lubans et al., 2012, p.2). Among their conclusions were:

- Physiological and psychosocial factors may explain the beneficial effects of physical activity programmes on social and emotional well-being in at-risk youth
- As none of the studies included long-term follow-ups (i.e., >12 months), it remains untested whether the benefits associated with participation in physical activity programmes are sustained once youth return to their daily routines

The reference to a lack of longitudinal research is also commented on by Jones et al. (2017), who found that only 5 percent of published USA sports studies included

a measure related to longer term youth development impacts. They suggest that analyses typically ‘end with short-term attitudinal outcomes which are often linked conceptually or theoretically with long term impacts, but seldom explored empirically’ (Jones et al., 2017, p. 15). In this regard, Morris, Sallybanks, Willis and Makkai (2003, p. 74) argue that after participation in programmes most participants will return to their previous environment and that; ‘a program on its own cannot effectively produce lasting changes in antisocial behaviour by young people—there is a need for continual care in the community that encourages maintenance of positive behavioural change.’ A major study of British rehabilitation programmes (Taylor, Crow, Irvine, & Nichols, 1999) concluded that evaluation was variable and that performance indicators ranged from the simple monitoring of attendance, via the use of anecdotal evidence to a few who estimated reconviction rates. Their conclusions illustrate many of our more general concerns with programme processes and mechanisms and methodological limitations and resulting policy dilemmas:

Programme managers...feel that quantitative indicators are insufficient to capture the essence of the outputs [and] that this reflects the difficulty of not only determining the significant variables but also measuring the precise effect they have....There is a problem finding qualitative evaluation techniques which are feasible with limited resources, but which adequately monitor the complex outcomes which most of the programmes aspire to. All programmes agree that physical activities do not by themselves reduce offending. All agree that there are personal and social development objectives that form part of a matrix of outcomes. These developments may, sooner or later, improve offending behaviour, but their impact is unpredictable in scale and timing. (Taylor et al., 1999, p. 50)

Another area in which a single variable intervention—sport—has been found to be problematic is the claims for the positive relationships between sports participation and improved educational performance. Grissom (2005) points to a widespread failure of experimental designs to find statistically significant differences between experimental and control subjects. He argues that this is due in part to the generic difficulty in raising academic achievement. It is very difficult to raise student achievement, beyond what might be expected, even when that is the specific focus. A study intended to affect achievement indirectly (i.e., via participation in PE and sport) encounters even more difficulty.

Shephard (1997, p. 115) illustrates the problems in assessing cause and effect as follows:

Even in studies where physically active students have had an unequivocal academic advantage over their sedentary peers, it is unclear whether intelligence led to success in sport, whether involvement in an ac-

tivity program enhanced academic performance, or whether both academic success and a predilection for physical activity are related to some third factor, such as a genetic characteristic that favors both academic and physical developments.

Grissom (2005) admits that the understanding of these issues will probably not be achieved via experimental or correlational designs and suggests that there is an urgent need for ‘naturalistic’ research to understand mechanisms and contribute to the building of theory.

3.3. Sufficient Conditions

The reference to the understanding of processes and mechanisms and how programmes are meant to work raises the third limitation of much existing published research—the failure to consider *sufficient conditions*. Participation in ‘sport’ (however defined) is a necessary, but not sufficient condition to obtain the supposed benefits. There is a lack of information about the various mechanisms, processes and experiences associated with participation. Patriksson (1995, p. 128) argues that:

Sport, like most activities, is not a priori good or bad, but has the potential of producing both positive and negative outcomes. Questions like ‘what conditions are necessary for sport to have beneficial outcomes?’ must be asked more often.

West and Crompton (2001), in a review of 21 North American outdoor recreation programmes aimed at reducing recidivism, found a widespread absence of clear statements of rationale and associated theory about the *presumed relationships* between participation, changed attitudes and changed behaviour (see also Collins, Henry, Houlihan, & Buller, 1999; Witt & Crompton, 1996). Bidle, Gorely and Stensel (2004, p. 689) in a review of school-based interventions to increase physical activity among young people, state ‘the extant literature did little to improve understanding of *what kinds of programmes or what aspects of programmes* bring about health gains or valued outcomes’.

Jones et al. (2017) found that only 18 percent of articles included information related to the logic or rationale of the programme under study. From this perspective, Jones et al. (2017, p. 14) raise a fundamental question about the limitations of much descriptive outcome-based research—‘without this information it is unclear if the evaluative criteria used by the researchers matched the programme model, or if the constructs being measured were an intended or unintended consequence of participation’.

Details of the ‘middle-range mechanisms’ (Pawson, 2006) of programmes are frequently missing, either for conceptual reasons—they are not part of the methodological approach to research—or due to the requirements of journals with their rather standardised ap-

proach to formats and limitations on word length. In this regard, Pawson (2001) notes the limitations of narrative reviews. The process of information extraction is necessarily selective and, as social interventions and associated processes are descriptively inexhaustible, few research reports, research reviews or academic articles (the main source of information) will contain all relevant information. In other words, this approach is almost wholly dependent on published material and the decisions of authors, editors and the funders of research about the significance of issues for inclusion.

This is not an issue confined to sports research. Thornton and Lee (2000) refer to a general ‘publication bias’ and Oakley et al. (2005, p. 12) report that because of ‘reporting deficiencies...applicability and generalisability were limited by scant information’. Commenting on the general condition of social science research and its contribution to policy making Davies (2004, p. 13) concluded that there is a very strong need for ‘more and better implementation studies that can identify the particular conditions under which successful implementation and delivery takes place, or fails to take place, as well as those conditions that are more generalisable’.

These concerns raise the issue of a widespread need to understand the *processes* of participation—the nature of participants’ experience and the programme *mechanisms* which explain any measured changes in values, attitudes or behaviour. This refers to the need to ‘de-centre’ (Crabbe, 2000) (or ‘de-mythologise’) sport to understand what sports, work for what subjects, in what conditions and why? This need is indicated by Coakley (1998) who views ‘sports as *sites* for socialisation experiences, not *causes* of socialisation outcomes’ and Hartmann (2003) who argues that ‘the success of any sports-based social intervention program is largely determined by the strength of its *non-sport components*’. For example, Sandford, Armour and Warmington (2006) concluded that, ‘*social relationships* experienced during involvement in physical activity programmes are the most significant factor in effecting behavioural change’. The need for such an understanding was emphasised by the conclusion of the literature review by the Sport for Development and Peace International Working Group (2007, p. 4) that ‘the evident benefits appear to be an indirect outcome of the context and social interaction that is possible in sport rather than a direct outcome of participating in sport’. Such perspectives reflect a growing realisation that sports on their own have difficulty in achieving the variety of desired outcomes. This is indicated by the growth of *sport plus* approaches, in which sports’ contributions are complemented and supported by a range of parallel initiatives (Coalter, 2007). In such circumstances sport is mostly a vitally important necessary, but not sufficient, condition.

4. Theory-Based Evaluation

The concern with programme mechanisms reflects a growing recognition of the need for theory-based eval-

uation (TBE) approaches. Weiss (1997, p. 520) argues that ‘the clearest call for theory based evaluation comes when prior evaluations show inconsistent results’—a situation which neatly sums up sports-related research.

The key idea of theory-based evaluation is that policy makers’ and programme providers’ beliefs and assumptions underpinning any intervention can be expressed in terms of a ‘programme theory’ or theory of change—a sequence of presumed causes and presumed effects (Weiss, 1997). Not only does this approach seek to describe actual mechanisms, but it ‘aims to surface the theoretical underpinnings of the program in advance and use the theories to help structure the evaluation’ (Weiss, 1997, p. 510). Weiss (1997, p. 510) argues that:

One of the main reasons for interest in TBE is the usual inability of even the most sophisticated experimental evaluation to explain what factors were responsible for the program’s success—or failure. Although evaluations based on random assignment to program and control groups give good estimates of program impact, they have little to say about *how or why* the impacts occurred.

The World Bank (2004, p. 10) argues that:

Theory-based evaluation...allows....an in-depth understanding of a working of the program or activity—the ‘program theory’ or ‘program logic’. In particular it need not assume simple linear cause-and-effect relationships....By mapping out the determining or causal factors judged important for success, and how they might interact, it can then be decided which steps should be monitored as the progress develops, to see how well they are in fact borne out. This allows the critical success factors to be identified.

Such an approach has major methodological implications as it seeks to specify the mechanisms by which change is, or is not, achieved and not simply identify the activities and characteristics associated with change. In such circumstances, Weiss (1997, p. 514) outlines the methodological implications as follows:

One of the hopes of the theories-of-change approach was to obviate control groups....Hope is that TBE can track the unfolding of events, step-by-step, and thus make causal attributions based on demonstrated links. If this were so, evaluation would not need randomized control groups to justify its claims about causality.

There is now a widespread acknowledgment of the need for an understanding of programme processes—the nature of participants’ experience and the *mechanisms* which explain any measured changes in values, attitudes or behaviour. We have limited understanding about what sports and sports’ *processes*, produce what

outcomes, for which *participants* and in what *circumstances*. Consequently, the programme theory approach does not offer the policy community a ‘best buy’ but ‘a tailored, “transferable theory”—(this programme theory works in these respects, for these subjects, in these kinds of situations)’ (Pawson, 2001, p. 4). As Weiss (1997, p. 518) explains ‘TBE...provides explanations—stories of means and ends—that communicate readily to policy makers and the public’.

In terms of the processes of evidence-based policy the attraction of theory-based approaches to evaluation is that they provide an opportunity to close the distance between academic research, policy makers and practitioners and to move beyond simple ‘political arithmetic’ and ‘partisan support’ and embrace an approach in which the ‘influence of research on policy occurs through the medium of ideas rather than of data’ (Pawson, 2006, p. 169). For example, Weiss (1980) argues that a theory-based approach entails a ‘conversation’ between researchers, policy-makers and practitioners. As Bailey et al. (2009, p.31) suggest:

One of the key tasks for researchers is to work with programme developers and sponsors to analyse the outcomes for which they are hoping. More importantly, the analysis reveals assumptions (and micro-assumptions) that have been made about the ways in which programme activities will lead to intended outcomes. A theory of change approach to evaluation argues that this clarification process is valuable for all parties, particularly in making explicit powerful assumptions that may or may not be widely shared, understood or agreed.

Experience suggests that we may need to look lower down the intervention hierarchy and possibly find a more receptive audience among programme planners and managers—those with direct responsibility for delivering the inflated and often intellectually incoherent promises of policy makers. Of course, it could be argued that, to expose such inflated promises, it is necessary to engage with those who formulate them. However, where policymaking is dominated by political processes of opportunism, persuasion, negotiation, partnership building, the seeking of organisational and interest-group advantage and deeply embedded professional repertoires, entry is often difficult and usually confined to the ‘chosen few.’

However, the above issues are of vital practical importance to sports policy, provision, programme design and management. The conclusion from many of the areas considered in the monitor is that there is a need for better understanding of the *processes* of sports participation (mostly sport plus programmes) which lead, or do not lead, to desired policy outcomes (Jones et al., 2017). Such evaluations would seek to understand if failure to achieve desired outcomes was because of an inherently faulty programme theory, or poor and inconsistent im-

plementation. These issues relate not simply to attempts to advance academic knowledge, but also to improve practice and perhaps most importantly of all, to understand the strengths and limitations of the claims that can be made for 'sport' and 'social inclusion'—important issues for policy and practice.

Acknowledgements

I would like to acknowledge the contribution of John Taylor (Stirling University) for his extensive contribution to the compilation of the Value of Sport Monitor.

Conflict of Interests

The author declares no conflict of interests.

References

- Bailey, R., Armour, K., Kirk, D., Jess, M., Pickup, I., Sandford, R., & BERA Physical Education and Sport Pedagogy Special Interest Group. (2009). The educational benefits claimed for physical education and school sport: An academic review. *Research Papers in Education, 24*(1), 1–27.
- Biddle, S. J., Gorely, T., & Stensel, D. J. (2004). Health-enhancing physical activity and sedentary behaviour in children and adolescents. *Journal of sports sciences, 22*(8), 679–701.
- Biddle, S. J., O'Connell, S., & Braithwaite, R. E. (2011). Sedentary behaviour interventions in young people: A meta-analysis. *British Journal of Sports Medicine, 45*(11), 937–942.
- Bloom, M. R., Grant, M. W., & Watt, D. (2005). *Strengthening Canada: The socio-economic benefits of sport participation in Canada*. Ottawa: Conference Board of Canada.
- Blunkett, D. (2000). Influence or irrelevance: can social science improve government. *Research Intelligence, 71*, 12–21.
- Cabinet Office. (1999). *Modernising government*. London: Cabinet Office.
- Caborn, R. (2003). Transcript of Richard Caborn's opening speech at CCPR 20. Retrieved from http://www.culture.gov.uk/Reference_library/Minister_Speeches/Richard_Caborn/Richard_Caborn_Speech03.htm
- Coakley, J. (1998). *Sport in society: Issues and controversies* (6th ed.). Boston, MA: McGraw Hill.
- Coalter, F. (2007). *A wider social role for sport: Who's keeping the score?* London: Routledge.
- Collins, M., Henry, I., Houlihan, B., & Buller, J. (1999). *Sport and social inclusion: A report to the Department of Culture, Media and Sport, Institute of Sport and Leisure Policy*. Loughborough: Loughborough University.
- Crabbe, T. (2000). A sporting chance? Using sport to tackle drug use and crime. *Drugs: Education, Prevention and Policy, 7*(4), 381–391.
- Davies, P. (2004, February). Is evidence-based government possible. In *Fourth Annual Campbell Collaboration Colloquium* (Vol. 19).
- Department of Culture, Media and Sport, & Strategy Unit. (2002). *Game plan: A strategy for delivering government's sport and physical activity objectives*. London: Cabinet Office.
- Etnier, J. L., Salazar, W., Landers, D. M., Petruzzello, S. J., Han, M., & Nowell, P. (1997). The influence of physical fitness and exercise upon cognitive functioning: A meta-analysis. *Journal of sport and Exercise Psychology, 19*(3), 249–277.
- Faulkner, G., Taylor, A., Ferrence, R., Munro, S., & Selby, P. (2006). Exercise science and the development of evidence-based practice: A "better practices" framework. *European Journal of Sport Science, 6*(2), 117–126.
- Glasner, P. E. (1977). *The sociology of secularisation: A critique of a concept*. London: Routledge & K. Paul.
- Grissom, J. B. (2005). Physical fitness and academic achievement. *Journal of Exercise Physiology Online, 8*(1), 11–25.
- Harris, K., & Adams, A. (2016). Power and discourse in the politics of evidence in sport for development. *Sport Management Review, 19*(2), 97–106.
- Hartmann, D. (2003). Theorising sport as social intervention: A view from the grassroots. *Quest, 55*, 1118–1140.
- Houlihan, B., & White, A. (2002). *The politics of sports development: Development of sport or development through sport?* London: Routledge.
- Jackson, N. W., Howes, F. S., Gupta, S., Doyle, J. L., & Waters, E. (2005). Interventions implemented through sporting organisations for increasing participation in sport. *The Cochrane Library*.
- Jones, G. J., Edwards, M. B., Bocarro, J. N., Bunds, K. S., & Smith, J. W. (2017). An integrative review of sport-based youth development literature. *Sport in Society, 20*(1), 161–179.
- Long, J., & Sanderson, I. (2001). The social benefits of sport: Where's the proof? In C. Gratton & I. Henry (Eds.), *Sport in the city* (pp. 187–203). London: Routledge.
- Lubans, D. R., Plotnikoff, R. C., & Lubans, N. J. (2012). Review: A systematic review of the impact of physical activity programmes on social and emotional well-being in at-risk youth. *Child and Adolescent Mental Health, 17*(1), 2–13.
- Morris, L., Sallybanks, J., Willis, K., & Makkai, T. (2003). *Sport, physical activity and antisocial behaviour in youth*. Australian Institute of Criminology.
- Oakley, A., Gough, D., Oliver, S., & Thomas, J. (2005). The politics of evidence and methodology: lessons from the EPPI-Centre. *Evidence & Policy: A Journal of Research, Debate and Practice, 1*(1), 5–32.
- Patriksson, M. (1995). Scientific review Part 2. In *The significance of sport for society. Health, socialisation, economy: A scientific review* (Prepared for the 8th Conference of European Ministers responsible for

- Sport, Lisbon, 17–18 May 1995). Strasbourg: Council of Europe Press.
- Pawson, R. (2001). *Evidence based policy: 1. in search of a method* (Working Paper 3). London: Queen Mary University of London.
- Pawson, R. (2006). *Evidence-based policy: A realist perspective*. London: Sage.
- President's Council on Physical Fitness and Sports. (2006). *Sports and character development*. (Research Digest Series, 7/1). Washington, DC: President's Council on Physical Fitness and Sports.
- Rowe, N. (2005). *The value of sport*. Paper presented at Raising the Stakes. The Fourth All Island Sports Development Conference, Naas Co Kildare, Ireland.
- Sallis, J. F., Prochaska, J. J., & Taylor, W. C. (2000). A review of correlates of physical activity of children and adolescents. *Medical Science of Sports and Exercise*, 32(5), 963–975.
- Sandford, R. A., Armour, K. M., & Warmington, P. C. (2006). Re-engaging disaffected youth through physical activity programmes. *British educational research journal*, 32(2), 251–271.
- Shephard, R. J. (1997). Curricular physical activity and academic performance. *Pediatric exercise science*, 9(2), 113–126.
- Sibley, B. A., & Etnier, J. L. (2003). The relationship between physical activity and cognition in children: A meta-analysis. *Pediatric Exercise Science*, 15(3), 243–256.
- Solesbury, W. (2001). *Evidence based policy: Whence it came and where it's going* (ESRC UK Centre for Evidence Based Policy and Practice, Working Paper 1). London: University of London.
- Sport for Development and Peace International Working Group. (2007). *Literature reviews on sport for development and peace*. Toronto: Right to Play.
- Taylor, P., Crow, I., Irvine, D., & Nichols, G. (1999). *Demanding physical activity programmes for young offenders under probation supervision*. London: Home Office.
- Thornton, A., & Lee, P. (2000). Publication bias in meta-analysis: Its causes and consequences. *Journal of clinical epidemiology*, 53(2), 207–216.
- Weiss, C. H. (1980). Knowledge creep and decision accretion. *Knowledge: Creation, Diffusion, Utilisation*, 1(3), 381–404.
- Weiss, C. H. (1993). Where politics and evaluation research meet. *Evaluation practice*, 14(1), 93–106.
- Weiss, C. H. (1997). How can theory-based evaluation make greater headway? *Evaluation Review*, 21(4), 501–524.
- Weiss, C. H., & Bucuvalas, M. J. (1980). Truth tests and utility tests: Decision-makers' frames of reference for social science research. *American Sociological Review*, 45, 302–313.
- West, S. T., & Crompton, J. L. (2001). A review of the impact of adventure programs on at-risk youth. *Journal of Park & Recreation Administration*, 19(2), 113–140.
- Witt, P. A., & Crompton, J. L. (Eds.). (1996). *Recreation programs that work for at-risk youth*. State College, PA: Venture Publishing.
- World Bank. (2004). *Monitoring and evaluation: Some tools, methods and approaches*. Washington, DC: World Bank.

About the Author



Fred Coalter is visiting Professor of Sports Policy at the Free University of Brussels (VUB) and Leeds Beckett University. His published work includes *A Wider Social Role for Sport: Who's Keeping the Score?* (Routledge, 2007) and *Sport-for-Development: What Game Are We Playing?* (Routledge, 2013). He was a member of the UK's Deputy Prime Minister's Neighbourhood Renewal Unit's Working Group on Sport and the Centre for Social Justice's Working Group on Sport and Social Regeneration. Currently, he is a member of the Scientific Advisory Board of the Swiss Academy of Development.