

Challenging the New Orthodoxy: A Critique of SPLISS and Variable-Oriented Approaches
to Comparing Sporting Nations

Abstract

Research Question: In recent years the comparative sport policy field has become dominated by the ‘SPLISS’ approach developed by De Bosscher and colleagues. While this approach has developed important insights into the statistical relationship between key groups of independent variables and indicators of elite sport policy success, nevertheless its attempts to identify and explain both statistical association and causal relationships have significant limitations.

The paper thus seeks to address the question of the nature of such strengths and limitations and their implications for theory, policy and practice.

Methods: As a review paper it develops a critical evaluation of claims made for the SPLISS approach to variable oriented comparative policy analysis.

Results: The paper identifies and focuses on the implications of six key problems for the SPLISS approach, namely: philosophical assumptions and causal variables; the black box problem; internal validity issues; non-equivalence and reliability; the neglect of agency; and misconceptions in the use of mixed methods.

Implications: The paper’s findings represent a challenge to the hegemony of this variable-oriented approach and they argue not for replacement or rejection of such an approach, but for recognition of its limitations, and of the opportunities for complementing it with case-driven, qualitative analysis generating causal accounts of policy outcomes.

Keywords: sport policy, comparative, methodology; elite sport, sporting nations, SPLISS

Challenging the New Orthodoxy: A Critique of SPLISS and Variable-Oriented Approaches to Comparing Sporting Nations

Approaches to understanding the factors which lead to international sporting success, and in particular Olympic medal success, have evolved over time, but have intensified in the last two decades (Andersen & Roglan, 2012; Bergsgard, Houlihan, Mangset, Nødland, & Rommetveldt, 2007; De Bosscher, Bingham, Shibli, Van Bottenburg, & De Knop, 2008; De Bosscher, De Knop, Van Bottenburg, & Shibli, 2006; De Bosscher, Shibli, Westerbeek, & Van Bottenberg, 2015; Digel, Barra, Bosel, Fahrner, Kohl, & Utz, 2002; Digel, Burk, & Fahrener, 2006). These studies might be categorised as: critical description of the development of elite sport systems; input-throughput-output studies which calibrate statistically the relationship between investment and sporting success; and meso-level theory and its application which seeks to understand the policy processes surrounding elite sport development.

Early comparative sports studies focused on the systems and practices adopted by the former Communist bloc, and in particular the Soviet Union and the German Democratic Republic to produce elite sport success. Accounts such as those by Riordan and others (Riordan, 1999; Riordan & Jones, 1999), constitute critical description of the philosophy and operationalisation of support for elite sport development systems highlighting particular issues such as talent identification and specialization at an early age, specialist sport schools, and investment in sports science support, with largely implicit reference to systematic doping strategies. This tradition of critical description is evident also in the analysis of national sports policies produced by Laurence Chalip and colleagues in their overview of sport policy systems in seventeen countries, with varying degrees of critique expressed by authors from the different contributing countries (Chalip, Johnson, & Stachura, 1996).

1
2
3 In an earlier paper on methodology in comparative sport policy studies, Henry,
4 Amara, Al-Tauqi, and Lee (2005) introduce a typology of four methodological approaches to
5 comparative analysis, namely statistical comparison ('Seeking Statistical Similarities'),
6 qualitative description ('Describing Differences'), embedding accounts of national policy in
7 the context of theoretical accounts of change in the transnational environment ('Theorising
8 the Transnational'), and the discursive construction of policy systems ('Defining Discourse').
9
10 Each of these approaches relies on different ontological and epistemological assumptions
11 (which are usually implicit) and thus engages in a comparative analysis of a significantly
12 different order, each with their own strengths and weaknesses. The approach of Chalip et al.
13 (1996) constitutes an example of 'Describing Differences', and can be contrasted with Digel
14 et al.'s (2002) analysis, which effectively seeks to establish comparison on the basis of
15 common data measures for a single sport (athletics), and constitutes an example of 'Seeking
16 Statistical Similarities'. There is also a range of studies in the literature which notes the
17 statistical relationship between macro factors such as GDP, population size, political system,
18 degree of urbanisation etc. which are not necessarily directly related to, even if they are
19 statistically associated with, elite sport success (De Bosscher, 2007).
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 In terms of studies that reflect statistical relationships between characteristics of
41 national context and elite sport performance, De Bosscher and her colleagues confirm these
42 earlier claims of the significance of GDP per capita and population size in explaining
43 variance in Olympic medal winning performance by employing stepwise regression analysis.
44 Their analysis indicates that, 52.4% of elite sport success can be attributed to population size,
45 national wealth and whether the nations' political system was part of the former communist
46 bloc (De Bosscher, Heyndels, De Knop, van Bottenburg, & Shibli, 2008, p. 225).
47
48
49
50
51
52
53
54
55

56 The authors go on to argue that the residuals for each country (which represent the
57 unexplained variance in the dependent variable of the regression analysis) incorporate the
58
59
60

1
2
3 impact of sports policies aimed at securing medal success – though it is not possible to say
4
5 what proportion of the unexplained variance which is represented in the residuals, is
6
7 attributable to such factors. Indeed, the 52.4% of variance explained by these macro factors
8
9 may already incorporate an element of the variance explained by policy factors. For example,
10
11 a variable such as GDP per capita may be related to elite sports policy variables such as the
12
13 level of financial investment in elite sport (richer countries are likely to have greater
14
15 ‘disposable income’ to invest in sport) but it is not the fact of national wealth which might
16
17 explain performance, but rather what proportion of national resources are allocated to the
18
19 elite sport system and how such investment is used, which will provide rational explanation
20
21 of why performance varies between nation states.
22
23
24

25
26 The study by Digel et al. (2002) is, in a sense, a forerunner of what the SPLISS
27
28 studies have taken forward in a more comprehensive fashion in the context of an explicit
29
30 input-throughput-output approach. The Digel study adopted both quantitative and qualitative
31
32 data analysis in its comparison of the elite level systems of track and field athletics in eight
33
34 countries and their account incorporates a relatively transparent summary of difficulties of
35
36 data collection and analysis for this type of study. These problems include the German-
37
38 centric assumptions of the team of researchers which limit both the nature of the questions to
39
40 be raised and the terms in which such questions are framed and understood, as well as the
41
42 lack of comparability of statistical data (Digel et al., 2002, p. 78).
43
44
45
46

47 The work of Green and Oakley (2001) was also a precursor to the SPLISS studies, in
48
49 that the authors sought to establish key success factors associated with international sporting
50
51 success, and in particular the transfer of policy learning in the production of elite sporting
52
53 success from the former eastern bloc states. However, this analysis, unlike the SPLISS
54
55 studies, represents a qualitative case-based analysis of broadly comparable countries, rather
56
57 than an analysis of variables statistically associated with elite sport performance outcomes.
58
59
60

1
2
3 Green and Houlihan (Green, 2004; Green & Houlihan, 2004, 2005; Houlihan, 2013;
4
5 Houlihan & Green, 2008) develop further explanations of the adoption of particular elite
6
7 sport policy approaches drawing on meso-level theories of policy learning and transfer,
8
9 including application of the Advocacy Coalition Framework (ACF) (Sabatier & Jenkins-
10
11 Smith, 1993), path dependency (Kay, 2005; Peters, Pierre, & King, 2005), mimetic
12
13 isomorphism (DiMaggio & Powell, 1983), and regime theory (Houlihan, 2009; Yamamoto,
14
15 2009). Similarly, Andersen and Ronglan (2012) and Bergsgard et al. (2007) adopt
16
17 institutional frameworks to examine the similarity and convergence of elite sport policies in
18
19 Nordic and western nations respectively. This tradition of meso-level analysis of policy may
20
21 be described as comparative analysis which represents what Henry et al. (2005) refer to as
22
23 the ‘Seeking Similarities’ approach to comparative analysis, though Yamamoto’s analysis of
24
25 anti-doping regimes may be described as ‘Theorising the Transnational’ since it seeks to
26
27 uncover global-local relations in the development and implementation of anti-doping policy.
28
29 Nevertheless, although such approaches provide insights into agenda setting, inputs,
30
31 implementation, and some aspects of impact, they do not generally address policy outcomes
32
33 or causal relationships between inputs, throughputs, outputs, and outcomes in the manner
34
35 explicitly adopted by the SPLISS studies.

36
37
38 What is unique about the SPLISS studies themselves is the attempt to generate
39
40 evidence for a causal relation between particular policies, or groups of policies, and elite
41
42 sport outcomes. The SPLISS studies adopt the rhetoric and the conceptual architecture of
43
44 input-throughput-output-outcome analysis but has evolved over time. This evolution is
45
46 evidenced in the two major waves of studies (SPLISS 1.0, and SPLISS 2.0) initiated by the
47
48 SPLISS team. The SPLISS 1.0 papers relate to a six country study involving data from
49
50 Belgium (separated into data for Flanders and Wallonia), Canada, Italy, the Netherlands,
51
52 Norway, and the United Kingdom (De Bosscher, et al., 2008; De Bosscher, De Knop, & van
53
54
55
56
57
58
59
60

1
2
3 Bottenburg, 2009; De Bosscher, De Knop, van Bottenburg, Shibli, & Bingham, 2009; De
4
5 Bosscher, Shibli, van Bottenburg, De Knop, & Truyens, 2010). SPLISS 2.0 (De Bosscher et
6
7 al., 2015) is based on a wider study of 15 nations, composed of three of the SPLISS 1.0
8
9 nations, namely Belgium (Wallonia and Flanders), Canada, the Netherlands, and 12 others
10
11 (Denmark, Estonia, Finland, France, Northern Ireland (UK), Portugal, Spain, Switzerland,
12
13 South Korea, Japan, Australia, and Brazil).
14
15

16
17 The broad strategy of SPLISS 1.0 is much more explicitly positivistic than the
18
19 approach adopted in the 2.0 version which sought to expand its approach through adopting
20
21 what its authors describe as a mixed methods approach, though we argue below that there are
22
23 difficulties with this claim. SPLISS 1.0 set out to find the unique route to sporting success,
24
25 ignoring the possibility of equifinality, that is of reaching the same endpoint in more than one
26
27 way. The authors acknowledge this issue in the introduction to version 2.0
28
29

30
31 Where we naively started the SPLISS 1.0 project thinking that we could
32
33 identify a uniform best practice pathway towards building a perfect elite
34
35 sport development system, we now know that it is not so much the whole
36
37 of the system structure, but much more the unique combination of system
38
39 pieces that result in a variety of different approaches that deliver elite sport
40
41 success (De Bosscher et al., 2015, p. 15).
42
43
44

45 However, though this aspect of the problem is acknowledged, that is, that there is no
46
47 single configuration which uniquely generates elite sport success (see De Bosscher et al.,
48
49 2015, p. 5), SPLISS 2.0 still pursues the goal of providing a one size (or one sample)
50
51 approach to the problem, suggesting that better answers will be found by dealing with a
52
53 bigger sample of countries, more highly developed instruments, and a larger number of
54
55 dedicated researchers. (*ibid*, p. 15). This approach misses the critical point that national elite
56
57 sports policy systems are not all drawn from the same population. Some will be from the top
58
59
60

1
2
3 echelon of medal winning nations for which the use of the number of medals won, or market
4
5 share of medals won, might be a broadly appropriate dependent variable or measure of elite
6
7 sporting success. Others will be drawn from the group of nations which, while not necessarily
8
9 competing on a wide front for medals, nevertheless specialize in particular sports (e.g.
10
11 Kenya/Ethiopia in distance running), while others still may have little or no record of medal-
12
13 winning, and thus for these nations other measures of elite success will have to be adopted
14
15 (e.g. performance at regional level competitions).
16
17
18

19
20 Notwithstanding the advances which have been made in terms of sophistication of
21
22 analysis, there are a number of underlying limitations with the SPLISS approach that remain.
23
24 These problems are probably best described as limitations, in the sense that they do not
25
26 invalidate the analysis premised upon the SPLISS approach, but rather they highlight the
27
28 limits of the claims which might be made on the basis of the SPLISS approach. The purpose
29
30 of this paper, therefore, is to offer a qualified critique of variable-orientated approaches to
31
32 comparative sport policy analysis in general and to identify a number of limitations of the
33
34 SPLISS approach specifically. Below we outline six interrelated limitations of the SPLISS
35
36 approach.
37
38
39
40

41 **Six Key Limitations of the SPLISS Approach**

42 **Problem 1: Philosophical Assumptions and the Identification of Causal Variables**

43
44 The key goal of the SPLISS studies is to identify the critical success factors explaining elite
45
46 sport success. This is operationalized by establishing the statistical relationship between a
47
48 dependent variable usually expressed as numbers of Olympic medals won (but with
49
50 consideration given to a range of alternatives, including points per medal; De Bosscher et al.,
51
52 2015, p. 87), market share of medals (Shibli, & Bingham, with Henry, 2007), losing Olympic
53
54 finalists / diplomas (De Bosscher et al., 2015, p. 101), and groups of independent variables
55
56
57
58
59
60

described as ‘pillars’. In effect, the approach for SPLISS 1.0 and 2.0 has been to seek to establish the ‘facts’ of the causes of elite sport outcomes by reference to statistical associations of aggregated scores of Critical Success Factors (CSFs) for each pillar (based on ‘facts’ and stakeholder perceptions). As De Bosscher and colleagues note, “the SPLISS study was driven by the need to compare a large amount of data to gain insights into the analytical relationships between policy (input-throughput) and success parameters (outputs)” (De Bosscher et al., 2015, p. 3).

The authors, however, claim that their approach is more than one which seeks to establish statistical association between variables, and the implication that knowing what inputs are made and are associated with particular outputs does not allow us to claim that we can identify causes of particular outcomes. Knowledge that particular inputs are *associated with* given outcomes is no substitute for understanding / explaining *how* inputs are used to achieve or *cause* particular outcomes. Theory-driven approaches of success in achieving policy outcomes offer greater heuristic potential for explanation of such outcomes. Realist evaluation (e.g. Pawson, 2006, 2013; Pawson & Tilley, 2004), for example, focuses on causal mechanisms and their relationships with the local social, cultural, political, economic and organisational context, using theories of social change and / or logic models (Cooksy, Gill, & Kelly, 2001; Jordan, 2010) to inform ways of testing of theory-informed causal accounts of such change. (Archer et al., 1998, Pawson, 2006: 19-20, Pawson, 2009)¹. The point that explanation of change cannot be achieved by statistical association alone is acknowledged by the advocates of SPLISS (see De Bosscher, 2006 and De Bosscher at al., 2010) but not successfully addressed.

¹ While Pawson’s position is realist he rejects the critical theory stance implied in Bhaskar’s critical realism (Archer et al., 1998) opting instead for what he terms a ‘middle range realism’ (Pawson, 2006: 19-20; Pawson, 2009).

Both SPLISS 1.0 and 2.0 seek to elicit qualitative data which can reveal aspects of the internal mechanisms at play within the elite sport policy system. However, the nature of the data collected, and the manner in which data were collected, analysed, and employed (and the lack of detail in terms of protocols for how this has been done) make this element of the evidence appear to be far from systematic and rigorous, and thus their contribution to explanation is unconvincing.

The development of causal accounts which explain the outcomes of medal success would benefit from a more rigorous and developed causal framework. Such frameworks have been advanced in the policy analysis field, in for example the realist approach of Pawson and Tilley (Pawson, 2006; Pawson & Tilley, 1997). There are a number of benefits to this approach. The realist approach characterizes social policy contexts as open systems, explanations of which are required to place a greater emphasis on contextual factors in explaining outcomes. The approach which is summarized in Pawson and Tilley's 'formula' $C+M=O$: Context (C) plus causal Mechanism (M) explain Outcomes (O), implies that the Critical Success factors which the SPLISS approach rests upon, may well operate differently in different contexts. Of course SPLISS acknowledges the importance of economic and political contexts (the importance for example of the former communist, state-controlled athlete production systems; or of the neo-liberal sports economies such as that of the United States; or state intervention along the lines of social welfare economies, such as those of the Scandinavian welfare states). However, while the SPLISS team acknowledges the significance of different configurations of variables in different contexts it nevertheless underplays in its own explanations of cultural differences (whether generic, or sporting). In short, the SPLISS approach is seeking the answer to 'what works' in terms of producing elite sport success, whereas realist analysis addresses the question of 'what works? for whom?'

1
2
3 under what circumstances?', developing a more nuanced account in contexts which are
4
5 diverse and changing, in other words which constitute open systems.
6
7
8
9

10 **Problem 2: The 'Black Box' Problem**

11
12
13 Related to the underdevelopment of causal accounts of policy change is the 'Black Box'
14
15 problem, where the correlation between measurable inputs of resources, and outputs in terms
16
17 of results may be measured but the policy processes that are responsible for turning inputs
18
19 into outputs are not amenable to observation or measurement and thus are hidden from view
20
21 (hence the allusion to the policy implementation system as a 'black box'). The SPLISS
22
23 authors are obviously aware of the 'Black Box' problem but claim to have resolved this.
24
25
26

27
28 The SPLISS study does not just identify 'what' characterizes successful
29
30 elite sport policies, but also 'how' these different dimensions can be
31
32 developed. The unique feature of the research is that in addition to
33
34 measuring easily quantifiable variables such as inputs (e.g. money) and
35
36 outputs (e.g. medals) it [SPLISS 1.0] has also tried to assess the 'black
37
38 box' of throughput both in terms of the existence of various system
39
40 components and also the rating that athletes, coaches and performance
41
42 directors gave to these system components (De Bosscher et al., 2008, p.
43
44 35).
45
46
47

48 The authors characterise their explanation as identifying inputs in pillar 1, that is largely the
49
50 input of financial resources, and the use of this funding for facilitating the development of the
51
52 remaining pillars which are conceptualized as throughputs. However, treating variables such
53
54 as the number of coaches qualified and employed at various levels (Pillar 7) is hardly a
55
56 throughput. It is not the number of coaches employed within (or input into) the system which
57
58 brings about success, but rather it is how such coaches are utilized or employ their skills
59
60

1
2
3 which explains the success of the system. Similarly, it is not the input of money into the
4
5 sports science support system which explains success, but how that support system is
6
7 organized to leverage success. The input is sports science knowledge and resources, the
8
9 throughput is the way in which sports science is employed, the nature of the causal
10
11 mechanisms identified and employed by those delivering the programme.
12
13

14
15 This issue is bound up with the nature of qualitative data in the specification of
16
17 throughputs. If we are to describe or define throughputs, we are implicitly bound to the
18
19 process of specifying (or at the very least speculating upon) causes of outcomes. However,
20
21 when we dig down into the ways in which qualitative data are obtained and analysed (as we
22
23 note under our discussion of the misuse of the term mixed methods in discussing our sixth
24
25 ‘problem’ below), there are significant limitations to the usefulness of this information.
26
27 Explaining what is happening within the policy system (what is going on within the ‘black
28
29 box’) in terms of factors critical to the success of the policy system implies more than simply
30
31 identifying the statistical association, i.e. the constant conjunction of input with output, or
32
33 even of throughput resources (which we conceive of as inputs) and outputs. Rather it involves
34
35 detailed qualitative accounts of how particular actions bring about particular outcomes,
36
37 requiring detailed *qualitative analysis* of the qualitative data obtained.
38
39
40
41
42
43
44
45

46 **Problem 3: Internal Validity**

47
48 Internal validity relates to the issue of whether outcomes observed in a study / analysis are
49
50 due to changes in an independent variable or variables, and not to some other, external,
51
52 factor(s). In other words, internal validity implies a causal relationship between the
53
54 independent and dependent variable. Internal validity is related to the extent to which an
55
56 operational measure captures the reality of the concept to be measured, and is said to exist if
57
58
59
60

1
2
3 the observed variation of the independent variable(s) affects the value of the dependent
4
5 variable (e.g. medals or market share).
6
7

8 The dependent variable for the SPLISS studies has tended to be operationalised in
9
10 terms such as Olympic medals won, or ‘market share’ of medals achieved. Given that by
11
12 2018 after the Pyeongchang Games, 34% of National Olympic Committees had never won a
13
14 medal in either Winter or Summer editions of the Olympics, of any sort, and 78% had never
15
16 won a medal in the Winter Games, it is clear that as a measure of ‘success’ in elite sport
17
18 terms, this is a relatively blunt instrument for a large proportion of nation states. Thus, a more
19
20 useful approach might be to cluster nations into appropriate performance brackets and to
21
22 evaluate nations by comparison with other nations with similar sporting profiles (e.g. using
23
24 dependent variables of performance in continental or regional games and / or using resource
25
26 capacities such as GDP to group cases together, thus ‘comparing apples with apples’), or to
27
28 identify success in terms of improvement of performance by self-benchmarking of nations
29
30 over time. A further complication relates to the treating of medals for all sports as being of
31
32 equivalent value which may not reflect national priorities, and thus may not be a culturally
33
34 appropriate, or valid measure of success.² Nations may also adopt more specific priorities
35
36 such as placing a value on achieving equity goals (e.g. having a relatively even spread of
37
38 medals across both genders in Olympic and / or Paralympic competition).
39
40
41
42
43

44 Perhaps the dominant view in relation to comparative analysis in the social sciences,
45
46 is that such analysis should be based on analysis at the level of variables. However, as Della
47
48 Porta (2008: 198) argues,
49

50
51 ... much research – especially in political science, but also in some

52
53 branches of sociology – is case-oriented: that is, it aims at rich descriptions
54
55

56
57
58
59 ² Medals for shooting, for example, may be less valued in some countries than medals in track and field
60 athletics, though some countries such as Kuwait and Malta clearly value success in shooting highly, since
shooting sports are central to the sporting cultures of those countries.

1
2
3 of a few instances of a certain phenomenon. ... Variable-oriented studies
4
5 mainly aim at establishing generalized relationships between variables,
6
7 while case-oriented research seeks to understand complex units.
8
9

10 As Della Porta suggests, both approaches are legitimate, but in the case of the SPLISS
11
12 analysis of individual, or groups of, variables, case level analysis is neglected. This is
13
14 particularly pertinent in, for example, failure to identify cases for which medals won is not a
15
16 valid or appropriate dependent variable.
17
18

19 An inherent danger of the variable-oriented approach is that whilst it may be
20
21 methodologically convenient to apply a 'one-size-fits-all' nine-pillar framework across all
22
23 sporting nations, it assumes similarities of cases, overlooking (or ignoring) fundamental
24
25 differences that may exist within sporting systems. Thus, the deductive application of a pre-
26
27 determined framework, ignoring the development of inductive insights into the differences
28
29 between cases, has implications for sampling equivalence (see discussion of limitation 4
30
31 below).
32
33

34
35 Furthermore, the assumption' often made by variable-orientated researchers seeking
36
37 large-N comparisons is that the greater the number of variables employed in the analysis, the
38
39 'closer' or 'better' a reflection of reality the data are assumed to provide (Landman &
40
41 Carvalho, 2017). The SPLISS account has sought to increase its size and scale. In SPLISS
42
43 2.0, for example, the authors highlight that "the project involved 15 nations and responses
44
45 from 3000 elite athletes, over 1300 coaches and 241 performance directors...in collaboration
46
47 between 58 researchers and 33 policymakers" (De Bosscher et al., 2015, pp. 21-22) and that
48
49 "the elite sport climate survey, [was] completed by 3142 athletes, 1376 coaches and 243
50
51 performance directors" (De Bosscher et al, 2015, p. 12). The implication here is that more
52
53 data implies greater explanatory power.
54
55
56
57
58
59
60

1
2
3 Whilst the SPLISS studies should be acknowledged as one of the largest empirical
4 studies of its kind within the sport management domain, the approach adopted of expanding
5 the numbers of nations in the sample implicitly assumes that increasing the number of nations
6 involved in the study will enhance (confidence in) its findings. As we have already indicated
7 the validity of medals as a measure of elite sporting success is weak because countries do not
8 all belong to the same population as sporting nations, with, in many cases, nations not
9 winning any medals at all. Thus different sub-populations of nations require different
10 measures of sporting success if these are to be valid. Qualitative analysis of what constitutes
11 valid measures of sporting success for individual or groups of nations (in particular those
12 which do not win Olympic medals) is likely to produce more telling insights than simply
13 increasing the numbers of variables employed in developing nomothetic models.
14
15
16
17
18
19
20
21
22
23
24
25
26
27

28 With regard to the creation of the nine-pillar framework itself, the SPLISS team
29 claims this to be a conceptual model of the sports policy factors leading to international
30 sporting success that was derived from a systematic review of existing literature (though this
31 would seem to relate to a literature review which was conducted rigorously, rather than to the
32 conducting of a Systematic Review *per se*: Petticrew and Roberts, 2005). An important
33 fundamental question remains as to whether the nine-pillar depiction derived by De Bosscher
34 and colleagues (De Bosscher et al., 2006; De Bosscher et al., 2009) is indeed a framework or
35 model. These words are often used inter-changeably by the research consortium. Such
36 arguments may be viewed as semantic but are important when establishing whether the
37 SPLISS account can claim a causal relationship between its pillars (independent variables)
38 and the outcome variable (e.g. medals success) or between the pillars themselves. This issue
39 is central to understanding the nature and purpose of the SPLISS account and whether it can
40 claim to be more than an elaborate descriptive account or an attempt to classify sport policy
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

factors. In other words, they are important in understanding whether SPLISS can claim to be predictive.

Problem 4: Non-equivalence and Reliability

Equivalence in essence refers to how we know that we are studying the same phenomena in different contexts; in other words, how we know that the concepts and instruments utilized measure the same things in different contexts (Dowling, Brown, Legg, & Grix, 2018). The issue of (non-)equivalence is a methodological challenge for all comparative researchers and it is important to recognize that equivalence issues are pervasive at all levels of comparative analysis (Øyen, 2004; Ebbinghaus, 2005; Dowling et al., 2018; Jowell, 1998). However, we would argue that it is particularly problematic for large-scale, variable-orientated comparative studies like SPLISS.

De Bosscher and colleagues assume construct, functional and sampling equivalence. Construct equivalence refers to ensuring the instruments employed measure the same variables across nations. This is particularly challenging for sport policy scholars and as taken-for-granted concepts (e.g. athlete, coach, elite, participation, sport) in one country or language, are not necessarily equivalent in another as Digel and colleagues (2002) discovered.

Another fundamental issue surrounding equivalence is the compatibility and comparability of the secondary data used within the SPLISS studies. The comparative methodology literature refers to this as functional equivalence i.e. whether the data collected is standardized and contextualized to allow for meaningful comparisons (e.g., Øyen, 2004; Ebbinghaus, 2005; Landman & Carvalho, 2017). Comparative studies even of financial data on sport, within relatively similar governmental and economic systems have manifested major problems in terms of compatibility of data and interpretation as attempts to provide

cross national comparison even within Europe have highlighted (Andreff, 1997; Andreff, with Bourg, Halba, & Nys, 1995; Jones, 1989). De Bosscher and her colleagues acknowledge this in their own study “...comparisons of sport expenditure are challenging as expenditure definitions and sport policy delivery mechanisms vary considerably from nation to nation” (De Bosscher et al., 2015, p. 109). The data relating to contribution of financial support (Pillar 1) are therefore at best an approximate measure of financial input calculated on the basis of heterogeneous definitions and operationalisations of financial investment in sport. Similar issues of functional comparability arise when attempting comparison of data relating to other pillars, such as pillar 3, participation.

Comparative research also requires consideration of sample equivalence (Øyen 1990; Ebbinghaus, 2005; Jowell, 1998; Schuster, 2007). The critical factor to consider here is whether national systems are alike or differ in significant ways. Most comparative sport policy studies to date have chosen to adopt a ‘Most Similar Systems Design’ (MSSD) (Ragin, 2006), which involves comparing key features that are different amongst similar countries while controlling for a dependent variable. This is partly due to scholars being interested in comparing their own country with others, but also because they recognize that fairly similar countries are more likely to be comparable since they might operate in similar types of context (Lijphart, 1971).

The selection of cases for inclusion in SPLISS 2.0 is based on a number of criteria, but broadly speaking inclusion criteria represent pragmatic considerations. For example, although there are a wider number of cases involved in SPLISS 2.0, the inclusion of new countries is based on the availability and willingness of researchers or teams within these countries who have access to relevant data relating to the elite sport system in their own country. In SPLISS 2.0, “any nation interested was invited to participate under the condition that they were able to collect the comprehensive data set and follow research protocol” (De

Bosscher et al., 2015, p. 66). Thus, the nations incorporated in the SPLISS 2.0 study do not include countries which have won no medals at all in Summer Olympics (35% of all nations competing) or Winter Olympics (78%) nor does it include the strongest countries in terms of overall performance. Australia is included and has the best medal performance per head and per GDP of all countries for the Summer Games, but GB is excluded, because, although it participated in SPLISS 1.0, it feared that its market position in terms of understanding Key Success Factors, might be eroded if competitor nations gained access to its own intelligence on such matters (though Northern Ireland, one of the Home Nations making up Great Britain and Northern Ireland, is included).

The decision to select a sample of nations on a pragmatic basis, whilst convenient, runs the risk of selection bias or 'problem of contingency' (Ebbinghaus, 2005). In the case of SPLISS and other large-N comparative studies 'selection bias' should be acknowledged and its implications accounted for.

The pragmatic nature of sample selection means that the SPLISS framework is unable to account for some of the unique system features of the 'powerhouse' medal-winning nations, such as, for example, the role of the education system in the US, or the centralized co-ordination role of the Chinese government in producing elite athletes. In addition, although there has been some shift with regard to less economically developed nations from SPLISS 1.0 to SPLISS 2.0, the SPLISS nation sample continues to focus predominantly on developed nations and hence the SPLISS account has far less utility in relation to less economically developed nations. Having a stratified sample is not of itself problematic, as such an approach to sampling might allow 'like-for-like' comparison between groups of nations. Clustering of nations into sub-samples with similar characteristics would allow of more meaningful and relevant analysis with the potential to generate more meaningful explanation of different types of performance.

1
2
3 A final set of problems, which are linked to equivalence, relates to the inter-observer
4 and study reliability of the SPLISS account. The reliance on local researchers in SPLISS 1.0
5 and 2.0 within each national context brings with it potential difficulties. Of course this is a
6 strategy for increasing the reliability of local data collection since as these researchers are
7 able to assist the consortium in ensuring, for example, construct equivalence of concepts
8 employed – some of which are evidently not appropriate and/or applicable in certain national
9 contexts. The SPLISS consortium seeks to control for such difficulties by providing a
10 protocol for data definition and collection. Nevertheless, the practical reality of this approach
11 to data collection is that the exercise is still carried out by an individual researcher or a small
12 team who are responsible for assessing their own nation's sporting system, including the
13 social and political environment, sporting history and culture. The researcher(s) is / are
14 required to translate their own observations and data into inventories of observations in a
15 centrally devised framework provided by the consortium. This process of standardization of
16 data leaves considerable leeway both for researcher bias, and for the suppression of cultural
17 specificity in individual sport policy systems.
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

38 Individual researchers and research teams participating in the SPLISS projects are
39 self-funding. This is a practical limitation in itself in that involvement in the project is
40 possible only for individuals with the necessary economic and social capital to be able to
41 invest the time, and to access appropriate data required. This practical resource limitation
42 may, in part, explain why the SPLISS nation samples are still dominated by western-northern
43 nations despite the increase in sample size between SPLISS 1.0 and 2.0. It is also important
44 to acknowledge the motivations of funding bodies and their limitations in relation to access to
45 data. For example, De Bosscher and colleagues note that perhaps the major motivation of
46 funders for participating in the SPLISS studies is to gain greater knowledge of critical
47 success factors in order to obtain a market advantage in sporting performance. However, the
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 public availability of data and findings from the SPLISS studies means that this market
4
5 intelligence is open to all, and where winning medals is a zero-sum game the national
6
7 sporting interests of participating nations may be damaged by the ability of their rivals,
8
9 whether or not they have funded a study as part of the SPLISS project, to access this
10
11 information. As De Bosscher and her colleagues note, this factor explains the withdrawal of
12
13 UK Sport's support for British involvement in SPLISS 2.0, having been a principal actor in
14
15 the SPLISS 1.0 consortium "many nations were looking at the UK as a best practice
16
17 benchmark and as such the UK felt less eager to take part in SPLISS 2.0" (De Bosscher et al.,
18
19 2015, p. 67).
20
21
22

23
24 De Bosscher and colleagues have attempted to mitigate the above-mentioned issues of
25
26 equivalence. Such mitigation strategies include utilizing European-wide secondary data sets
27
28 (e.g., Eurobarometer survey), data triangulation, generating detailed work protocols, defining
29
30 key terms/concepts, and translating their survey instruments into multiple languages.
31
32 Nonetheless, the employment of these strategies, while welcome in *mitigating* the effects of
33
34 such (non-)equivalence issues, does not mean that such problems have no negative impact on
35
36 inter-nation comparison (Jowell, 1998).
37
38
39
40
41
42

43 **Problem 5: Overlooking/Ignoring Agency**

44
45 The selection of the unit of analysis is integral to any social research endeavour, but is
46
47 particularly important for comparative researchers. Selection of the unit of analysis, the major
48
49 entity being studied, is fundamentally rooted in the researchers' ontology and epistemology,
50
51 as selection largely depends upon what sorts of knowledge are sought, and thus what
52
53 methodological approaches can be used to 'gather' or 'construct' such knowledge. It is here
54
55 that De Bosscher and colleagues take a particularly clear-cut stance in that they focus upon
56
57 the meso-level arguing that macro-level factors "are relatively stable and cannot be
58
59
60

1
2
3 influenced significantly by direct human influence such as policymakers” (De Bosscher et al.,
4
5 2015, p. 37). We would argue, contrary to De Bosscher and colleagues, that it is not
6
7 analytically possible to separate the meso from the macro or micro level. Rather, we adopt a
8
9 view more closely aligned to what Ragin (2014) refers to as a *comparativist* approach in that
10
11 we believe it is important to engage with and define macro-entities which influence the
12
13 policy process. In contrast, the approach taken by De Bosscher and colleagues can be
14
15 described as a non-comparativist, in that they deliberately choose to ignore such broader
16
17 notions as abstractions that need not to be operationalized. We therefore agree with the
18
19 remarks of Jowell (1998) who states that comparative researchers “...should pay as much
20
21 attention to the choice and compilation of aggregate-level contextual variables as they do to
22
23 individual-level dependent and independent variables” (p. 197) and that “all too often,
24
25 analysts seem to compare national data sets in vacuo” (ibid).
26
27
28
29

30
31 The inherent danger of the non-comparativist approach is an over-emphasis on
32
33 structuralist explanations of elite sport policy, which simultaneously overlooks/ignores the
34
35 influence and impact of broader social, political, and cultural (macro-level) factors on the one
36
37 hand and does not account for the role of individual actors (micro-level) on the other. This is
38
39 perhaps unsurprising given that the adoption of a variable-orientated approach results in the
40
41 identification of variables (critical success factors) which by their nature are structural. It is
42
43 these structural variables that are then correlated to derive empirical generalizations about
44
45 policy domains and it is these generalizations which become explanations of elite sport
46
47 success. The explanation put forward is therefore one of resources i.e. the ‘have’ and ‘have
48
49 not’ nations, rather than necessarily one of competency i.e. how effectively such resources
50
51 are utilized.
52
53
54
55

56 Furthermore, the SPLISS account also assumes rationality within the policy process
57
58 in that policy is evidence-based (although this may quite often not be the case) and that the
59
60

1
2
3 identification of these key variables (i.e. pillars) and the relative 'performance' of one nation
4
5 against another is therefore likely to lead to policy-makers acting upon such findings seeking
6
7 to improve performance in these policy domains.
8
9

10 11 12 13 **Problem 6: Reductionism and Mixed Methods** 14

15
16 The SPLISS studies claim a mixed-methods approach in that they utilize both qualitative and
17
18 quantitative data collected through an inventory and stakeholder surveys. In the SPLISS 2.0
19
20 study, the former (inventory) involved 96 CSFs divided into 212 open-ended and closed
21
22 questions across the nine pillars. The latter (survey), involved primary quantitative measures
23
24 on a 1-5 Likert scale. These instruments are used to generate composite indicators and critical
25
26 success factors through a mixture of qualitative and quantitative data. We take issue with the
27
28 use of composite indicators (CI), which is a quantitative or qualitative measure derived from
29
30 a series of observed 'facts' (Freudenberg, 2003). In a more recent paper, De Bosscher (2018)
31
32 acknowledge a number of problems with adopting a composite indicator methodology to
33
34 compare elite sport policies. However, while there may be a rationale for developing
35
36 quantitative composite indicators based on qualitative data in quantitative research, such a
37
38 strategy should not be confused with developing a mixed methods approach. Mixed methods
39
40 approaches do not develop qualitative methods and subsequently translate data thus obtained
41
42 into quantitative measures. A mixed methods approach implies the subjecting of qualitative
43
44 data to *qualitative forms of analysis* (e.g. subjecting documentary or interview data to
45
46 qualitative forms of analysis such as thematic analysis, analysis of discourse, or ethnographic
47
48 content analysis).
49
50
51
52
53

54
55 The subjecting of qualitative data to quantitative analysis will almost invariably
56
57 involve a significant loss of meaning and thus of explanatory power, and if it does not, the
58
59 value of collecting such qualitative data at all, is likely to be questionable. SPLISS 2.0, for
60

1
2
3 example, incorporates a large number of interviews with key stakeholders, e.g. athletes,
4
5 performance directors, coaches etc. Much of the data produced was recorded in a manner
6
7 amenable to quantitative analysis (e.g. the question “do athletes receive nutritional advice?”
8
9 generated a response which was reported as a simple dyadic/binary (yes/no) response) and
10
11 that the potential for contributing to understanding was much diminished. This type of
12
13 approach can hardly be said to be the product of mixed method inquiry.
14
15

16
17 Rather than invalidating the SPLISS account, we see qualitative analysis as a
18
19 particularly promising area that has the potential to significantly enhance the SPLISS account
20
21 of elite sport. Not only do both qualitative and quantitative data provide complementary
22
23 insights into the policy process in general but the amount of effort that De Bosscher and
24
25 colleagues have expended on collection of qualitative data might be expected, given
26
27 qualitative, rather than quantitative, analysis of the data to generate some real insights into
28
29 policy throughputs in the development of elite sport policies both within and across nations.
30
31
32
33
34
35

36 **Conclusions**

37
38
39 The critique of the SPLISS approach to analysing the success of elite sports policies which
40
41 we have sought to develop in this paper, does not seek to deny the contribution to policy
42
43 explanation and evaluation which variable-based approaches to analysis, including SPLISS,
44
45 can make. However, what we have looked to identify are the limits of such analysis in
46
47 generating explanations of policy success and the lacunae in such explanations. Our claim is
48
49 that SPLISS, as an analytic or conceptual framework, is useful but that its application as a
50
51 model (for example in terms of prediction of performance) is at best limited and at worst
52
53 inappropriate.
54
55

56
57 We have noted difficulties associated both (a) with how the SPLISS approach has
58
59 been applied; and (b) with the nature of the SPLISS approach per se. In relation to the former,
60

1
2
3 the adoption of pragmatic or convenience sampling is, for example, fraught with difficulties
4
5 since this requires a comparison of sporting nations of very different histories, cultures,
6
7 resources and perspectives on sport, such that comparison is in principle undermined. In
8
9 particular in analysis of elite sport policy success, if some measure of medal winning is the
10
11 basis of constructing the dependent variable as a marker of success, then the efforts made to
12
13 effect success by very many nations which have had no medal success at all, and /or whose
14
15 sporting targets are reduced to a level which is more realistic for them, would render this type
16
17 of dependent variable an insensitive measure of elite sport policy success. In reality, the
18
19 world's sporting nations do not constitute a sufficiently homogeneous population to foster
20
21 meaningful analysis for all nations, and if meaningful generalisations are to be made, a more
22
23 helpful approach would be to undertake some form of cluster analysis, grouping together like,
24
25 or akin, nations (in sporting and in broader terms) and undertaking separate analyses for each
26
27 cluster.
28
29
30
31

32
33 In relation to the appropriateness of the SPLISS approach per se, the potential of the
34
35 approach to produce explanation of the causal links between independent and dependent
36
37 variables is limited. The attribution of causes of policy outcomes cannot rely solely on
38
39 statistical analysis of quantitative data relating in particular to inputs, throughputs and
40
41 outputs. Detailed qualitative accounts of policy development and implementation will
42
43 potentially provide explanations of particular actors', or groups of stakeholders', perspectives
44
45 and explanations of how, why, and under what circumstances, particular policy actions have
46
47 (or have not) proved effective in achieving desired policy outcomes. Evaluating the validity,
48
49 or more accurately for some qualitative researchers (and in particular discourse analysts;
50
51 Wood & Kroger, 2000) the *warrantability* of such claims, will allow us to identify and assess
52
53 explanations of causes of policy change in groups of nations. In addition, it will allow us to
54
55 identify and evaluate explanations of policy success within a given case or cases (i.e. within a
56
57
58
59
60

national system or systems). Such explanations may be unique to particular contexts, and thus will address at the individual case level (or even within the case at the individual sport level) the question of what works, for whom, under what circumstances, for how long etc. in line with the aspirations of realist policy analysis (Pawson, 2006). This argument is consistent with the approaches of other comparative methodologists (e.g., Øyen, 1990, Jowell, 1998), who have argued that “comparative research may have to shift its emphasis from seeking uniformity among variety to studying the preservation of enclaves of uniqueness among growing homogeneity and uniformity” (Sztompza, 1988, p. 215). Analysis at the level of individual cases can also help to address issues of equifinality in which different policy configurations achieve similar outcomes.

While we recognise the attempts made, particularly in the shift from SPLISS 1.0 to 2.0, to develop a mixed methods approach, we would suggest that the application of the term ‘mixed methods’ to the approach adopted is mistaken. What in essence is undertaken in the SPLISS usage of the ‘qualitative’ data, is a quantitative representation of the qualitative data which is then subject to quantitative analysis. For a truly mixed methods approach to have been developed, qualitative data, rather than being converted to a quantitative format, would have been subjected to qualitative analysis. This, however, is only possible where treatment of qualitative data, whether it is in within-case or across-case analysis, is more clearly tied to formal protocols for extraction of themes, development of support for arguments, or identification of the nature of discourses etc. Such data will be particularly helpful in unpacking explanations of throughput in which the nature of the relationship between context, causal mechanisms and outputs is dynamic, given the open systems nature of policy relating to social contexts. The use of formal protocols (thematic analysis, discourse analysis, ethnographic analysis etc) will assist not simply in the generation of types of explanation, but

1
2
3 also in relation to claims about how consistent, coherent or exhaustive, such explanations are
4
5 in relation to the matters under evaluation.
6
7

8 The hegemony of the SPLISS approach in the field of analysis of the factors critical to
9
10 elite sport policy success has been very evident in the mobilisation of an active ‘industry’ of
11
12 SPLISS papers and associated conferences and workshops. Alternative, and in particular
13
14 qualitative, forms of comparative analysis in the sport field have subsequently been
15
16 neglected. Our paper has sought to highlight the nature of the insights that the SPLISS
17
18 approach has already generated, or has the potential to generate, but also by implication to
19
20 highlight what the SPLISS approach is not able to tell us, and therefore what other forms of
21
22 analysis might promise, with a view to addressing the imbalance in the current literature.
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- Andersen, S., & Roglan, L. (2012). *Elite Sport: Same Ambitions, Different Tracks*. Copenhagen: Copenhagen Business School.
- Andreff, W. (1997). *Sports Financing in Europe - Towards a Transformation in the 21st Century*. In (Vol. News Articles Section, Play the Game Website). Copenhagen: Play the Game.
- Andreff, W., with Bourg, J. F., Halba, B., & Nys, J. F. (1995). *Les enjeux économiques du sport en Europe: financement et impact économique*. [The economic challenges of sport in Europe: finance and economic impact.] Paris: Editions Dalloz.
- Archer, M., Bhaskar, R., Collier, A., Lawson, T. & Norrie, A. (eds.) 1998. *Critical Realism: essential readings*, Abingdon, Oxon: Routledge.
- Bergsgard, N. A., Houlihan, B. M. J., Mangset, P., Nødland, S. I., & Rommetveldt, H. (2007). *Sport Policy: a comparative analysis of stability and change*. Oxford: Butterworth-Heinemann.
- Busetti, S., & Dente, B. (2017). Using process tracing to evaluate unique events: The case of EXPO Milano 2015. *Evaluation*, 23(3), 256-273.
- Chalip, L., Johnson, A., & Stachura, L. (Eds.). (1996). *National Sports Policies: an International Handbook*. Westport, Conn.: Greenwood Press.
- Chen, S., & Henry, I. (2019). Assessing Olympic legacy claims: Evaluating explanations of causal mechanisms and policy outcomes *Evaluation*.
doi:<https://doi.org/10.1177/1356389019836675>
- Cooksy, L. J., Gill, P., & Kelly, P. A. (2001). The program logic model as an integrative framework for a multimethod evaluation. *Evaluation and Program Planning*, 24(2), 119-128. doi:10.1016/s0149-7189(01)00003-9

- 1
2
3 De Bosscher, V. (2007). *Sports Policy Factors Leading to International Sporting Success*,
4
5 Brussels: VUB Press.
6
7
8 De Bosscher, V. (2018). A mixed methods approach to compare elite sport policies of
9
10 nations. A critical reflection on the use of composite indicators in the SPLISS study.
11
12 *Sport and Society*, 21(2), 331-355.
13
14
15 De Bosscher, V., Bingham, J., Shibli, S., Van Bottenburg, M., & De Knop, P. (2008). *The*
16
17 *Global Sporting Arms Race: an International Study on Sports Policy Factors Leading*
18
19 *to International Sporting Success*. Oxford: Meyer & Meyer.
20
21
22 De Bosscher, V., De Knop, P., & van Bottenburg, M. (2009). An analysis of homogeneity
23
24 and heterogeneity of elite sports systems in six nations. *International Journal of*
25
26 *Sports Marketing & Sponsorship*, 10(2), 111-131.
27
28
29 De Bosscher, V., De Knop, P., Van Bottenburg, M., & Shibli, S. (2006). A conceptual
30
31 Framework for Analysing Sports Policy Factors Leading to International Sporting
32
33 Success. *European Sport Management Quarterly*, 6(2), 185-215.
34
35
36 De Bosscher, V., De Knop, P., van Bottenburg, M., Shibli, S., & Bingham, J. (2009).
37
38 Explaining international sporting success: An international comparison of elite sport
39
40 systems and policies in six countries. *Sport Management Review*, 12, 113–136.
41
42
43 De Bosscher, V., Heyndels, B., De Knop, P., van Bottenburg, M., & Shibli, S. (2008). The
44
45 paradox of measuring success of nations in elite sport. *Belgeo: Revue Belge de*
46
47 *Geographie*, 2, 217-234.
48
49
50 De Bosscher, V., Shibli, S., van Bottenburg, M., De Knop, P., & Truyens, J. (2010).
51
52 Developing a Method for Comparing the Elite Sport Systems and Policies of Nations:
53
54 A Mixed Research Methods Approach. *Journal Of Sport Management*, 24(5), 567-
55
56 600.
57
58
59
60

- 1
2
3 De Bosscher, V., Shibli, S., Westerbeek, H., & Van Bottenberg, M. (2015). *Successful Elite*
4
5 *Sport Policies: an International Comparison of the Sports Policy Factors Leading to*
6
7 *International Sporting Success (SPLISS 2.0) in 15 Nations*. Maidenhead: Meyer and
8
9 Meyer
10
11
12 Della Porta, D. (2008). Comparative analysis: Case-oriented versus variable-oriented
13
14 research. In D. Della Porta & M. Keating (Eds.), *Approaches and Methodologies in*
15
16 *the Social Sciences: A Pluralist Perspective* (pp. 198-222). Cambridge: Cambridge
17
18 University Press.
19
20
21 Digel, H., Barra, M., Bosel, J., Fahrner, M., Kohl, M., & Utz, A. (2002). *Organisation of*
22
23 *High Performance Athletics in Selected Countries: Final report for the International*
24
25 *Athletics Federation*. Tübingen: University of Tübingen
26
27
28 Digel, H., Burk, V., & Fahrner, M. (2006). *High Performance Sport: an International*
29
30 *Comparison*. Tübingen: Brauer.
31
32
33 DiMaggio, P., & Powell, W. (1983). The iron cage revisited: Institutional isomorphism and
34
35 collective rationality in organizational fields. *American Sociological Review*, (2016)
36
37
38 Dowling, M, Legg, D, & Brown, P. (2018). Cross-Comparative Sport Policy Analysis and
39
40 Paralympic Sport. In Brittain, I., & Beacom, A. (Eds.), *The Routledge Handbook of*
41
42 *the Paralympic Games* (pp. 249–272). London: Routledge
43
44
45 Dowling, M, Brown, P. Legg, D, & Grix, J (2018). Deconstructing comparative sport policy
46
47 analysis: assumptions, challenges, and new directions. *International Journal of Sport*
48
49 *Policy and Politics*. 10:4, 687-704
50
51
52 Ebbinghaus, B., 2005. When less is more: selection problems in large-N and small-N cross-
53
54 national comparisons. *International Sociology*, 20(2), 133–152.
55
56
57 Freudenberg, M. (2003). Composite Indicators of Country Performance: A Critical
58
59 Assessment. *OECD, Science Technology and Industry Working Papers* 16: 1–24.
60

- 1
2
3 Green, M. (2004). Power, policy and political priorities: Elite sport development in Canada
4 and the United Kingdom. *Sociology of Sport Journal*, 21(4), 376-396.
5
6
7 Green, M., & Houlihan, B. (2004). Advocacy Coalitions and Elite sport policy change in
8 Canada and the United Kingdom. *International Review for the Sociology of Sport*,
9 39(4), 387-403.
10
11
12 Green, M., & Houlihan, B. (2005). *Elite sport development: Policy learning and political*
13 *priorities*. London: Routledge.
14
15
16 Green, M., & Oakley, B. (2001). Elite sport development systems and playing to win:
17 uniformity and diversity in international approaches. *Leisure Studies* (20), 247-267.
18
19
20 Henry, I., Amara, M., Al-Tauqi, M., & Lee, P. C. (2005). A typology of approaches to
21 comparative analysis of sports policy. *Journal of Sport Management*, 19(4), 480-496.
22
23
24 Houlihan, B. (2009). Mechanisms of international influence on elite sport policy.
25 *International Journal of Sport Policy*, 1(51-70).
26
27
28 Houlihan, B. (2013). Theorising the analysis of sport policy. In I. Henry & L.-M. Ko (Eds.),
29 *Handbook of Sport Policy* (pp. 11-22) London: Routledge.
30
31
32 Houlihan, B., & Green, M. (2008). *Comparative elite sport development : systems, structures*
33 *and public policy*. Amsterdam: Elsevier ; Oxford : Butterworth-Heinemann.
34
35
36 Jones, H. G. e. (1989). *The Economic Impact and Importance of Sport in Europe: a*
37 *European Study*. Strasbourg: Council of Europe
38
39
40 Jordan, G. B. (2010). A theory-based logic model for innovation policy and evaluation.
41 *Research Evaluation*, 19(4), 263-273. doi:10.3152/095820210x12827366906445
42
43
44 Jowell, R. (1998). How comparative is comparative research?. *American Behavioral*
45 *Scientist*, 42(2), 168-177.
46
47
48 Kay, A. (2005). A critique of the use of path dependency in policy studies. *Public*
49 *Administration & Society*, 83, 553-571.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 Landman, T. & Carvalho, E. (2017). *Issues and methods in comparative politics* (4th edition).
4
5 London: Routledge.
6
7
8 Lijphart, A. (1971). Comparative politics and the comparative method. *American Political*
9
10 *Science Review*, 65(3), 682-693.
11
12 Oakley, B., & Green, M. (2001). The production of Olympic champions: International
13
14 perspectives on elite sports development systems. *European Journal for Sport*
15
16 *Management*, 8(1), 83-105.
17
18
19 Øyen, E. (1990). *Comparative methodology: theory and practice in international social*
20
21 *research*. London: Sage
22
23
24 Pawson, R. (2006). *Evidence-based Policy: a Realist Perspective*. London: Sage.
25
26 Pawson, R. (2013). *The Science of Evaluation: A Realist Manifesto*. Leeds: University of
27
28 Leeds.
29
30
31 Pawson, R. 2009. Middle Range Realism. *European Journal of Sociology / Archives*
32
33 *Européennes de Sociologie*, 41, 283-325.
34
35
36 Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: SAGE.
37
38 Pawson, R., & Tilley, N. (2004). Realist Evaluation. *Evaluation*, 48, 1-36.
39
40
41 Pawson, R. D., & Boaz, A. (2004). *Evidence-based policy, theory-based synthesis, practice-*
42
43 *based reviews*. Swindon: Economic and Social Research Council.
44
45
46 Peters, B. G., Pierre, J., & King, D. S. (2005). The politics of path dependency: Political
47
48 conflict in historical institutionalism. *Journal of Politics*, 67, 1275-1300.
49
50
51 Petticrew, M & Roberts, H. (2005). *Systematic Reviews in the Social Sciences: A Practical*
52
53 *Guide*, Oxford: Wiley-Blackwell.
54
55
56 Ragin, C., (2006). How to lure analytic social science out of the doldrums: Some lessons
57
58 from comparative research. *International Sociology*, 21 (5), 633–646.
59
60

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- Ragin, C. (2014). *The comparative method: Moving beyond qualitative and quantitative strategies*. Oakland: University of California Press.
- Riordan, J. (1999). The Impact of Communism on Sport. . In J. Riordan & A. Kruger (Eds.), *The International Politics of Sport in the 20th Century* (pp. 48-66). London: Taylor and Francis.
- Riordan, J. (Ed.) (1978). *Sport under Communism*. London: Hurst.
- Riordan, J., & Jones, R. (1999). *Sport and Physical Education in China*. London: Taylor and Francis.
- Sabatier, P. A., & Jenkins-Smith, H. C. (1993). (Eds.), *Policy change and learning : an advocacy coalition approach*, Boulder, Colo. ; Oxford: Westview Press.
- Schuster, J. M. (2007). Participation studies and cross-national comparison: proliferation, prudence, and possibility. *Cultural trends*, 16(2), 99-196.
- Shibli, S., & Bingham, J., with Henry, I. (2007). Measuring the Sporting Success of Nations. In I. Henry & Institute of Sport and Leisure Policy (Eds.), *Transnational and Comparative Research in Sport: Globalisation, Governance and Sport Policy* (pp. 61-81). London: Routledge.
- Sztompka, P. (1988). Conceptual frameworks in comparative inquiry: divergent or convergent?. *International Sociology*, 3(3), 207-218.
- Wood, L., & Kroger, R. (2000). *Doing Discourse Analysis: methods for studying action in talk and text*. London: Sage.
- Yamamoto, M. (2009). *The influence of non-domestic factors on elite sport development and anti-doping policy: The cases of Japan and UK/England (Unpublished PhD thesis)*. Loughborough University, Loughborough.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review Only