

Indicators of good governance in sport organisations

Handle with care

Arnout Geeraert

Introduction

Having authored the Sports Governance Observer 2015 and 2018 (Geeraert 2015, 2018a), the National Sports Governance Observer (NSGO) (Geeraert 2018b) and the National Anti-Doping Governance Observer (Geeraert 2021), I am happy to see that these sets of indicators are increasingly applied across different fields to benchmark good governance in international sport federations, national sport federations, and national anti-doping organisations, respectively. Yet, I can state with certainty that they are all flawed: none of them provide a completely valid and reliable quantification of good governance. I do not think, however, that this is problematical per se. No measurement of governance is or will ever be perfectly reliable and valid. I do feel that these (inherent) limits are not always well-understood, particularly in the field of sport, where indicators that aim to assist and inspire good governance have been mushrooming in recent years without much critical reflection. As a result, indicators of good governance in sport organisations, including those I developed, risk inducing seriously misguided good governance policies.

The purpose of this chapter is therefore to question and challenge the development and use of indicators of good governance in sport organisations. Adding to the aforementioned instruments, a few other influential initiatives should be mentioned. Chappelet and Mrkonjic (2013) developed the 'Basic Principles of Governance in International Sport' (BIBGIS), the first academic attempt to formulate a comprehensive set of indicators for assessing good governance in international sport federations. The Olympic and sport movement promotes three particular sets of indicators. In 2008, the International Olympic Committee (IOC) introduced the 'Basic Universal Principles of Good Governance of the Olympic and Sports Movement' (PGG) (IOC 2008). In 2016, the Support the Implementation of Good Governance in Sport (SIGGS) project funded by the European Union (EU) Erasmus + programme and managed by the European Olympic Committees European Union Office introduced a set of indicators that built on a previous EU funded project (Zintz & Gérard 2019). They were developed to assist national Olympic committees and national sport federations

with the implementation of good governance through self-evaluation. Finally, the Key Governance Principles and Basic Indicators (KGP) were introduced by the Association of Summer Olympic International Federations in the same year (ASOIF 2016). The KGP are part of a voluntary monitoring system implemented and managed by ASOIF to achieve better governance in Olympic international sport federations (Geeraert 2019). In an increasing number of countries, including Belgium (Flanders), the Netherlands, the United Kingdom, and Australia, codes and derived indicators are applied by public or sport authorities to induce sport federations' compliance with standards of good governance.

The popularity of indicators of good governance in sport organisations can be explained by their potential for knowledge building. They inform policy formulation and decision-making and facilitate advocacy by civil society organisations. They do so by providing readily understandable information about performance and institutional quality and by alerting policymakers to specific trends and significant shortcomings (Coppedge et al. 2011; Espeland & Sauder 2007; Marlier & Atkinson 2010; Saisana & Tarantola 2002; Saltelli 2007). This is especially useful considering the uncertainty about the status quo of quality of governance in different types of sport organisations. From an academic point of view, moreover, gathering comparative governance data constitutes a necessary step towards testing theories about both factors that contribute to good governance and the impact of implementing good governance (Coppedge et al. 2011; Gisselquist 2012).

Despite these advantages, the increasing use of governance indicators in sport also entails substantial risks. While governance indicators are often perceived as certain and objective, in reality they are not. As Sarfaty (2015, p. 123) aptly observes, "behind a veil of scientific truth and neutrality, they mask potential problems". These problems can be quite severe. Indicators produce powerful knowledge effects—they increase awareness and induce acceptance of specific standards and practices, as well as governance effects—they influence decision-making and power dynamics (Merry 2016, pp. 4–5). Yet governance indicators tend to lack adequate methodological transparency, which leads to "questionable knowledge and governance effects" (Nelken 2015, p. 332). Most importantly, however, indicators that are poorly constructed, misinterpreted, manipulated or misused can produce flawed knowledge and misguided decisions (Malito, Umbach & Bhuta 2018; Nardo et al. 2005; Saisana & Tarantola 2002). In other words, indicators of good governance might just as well induce bad governance.

What exactly are the risks inherent to the use of governance indicators in sport and how, if at all, can they be dealt with? In answering these questions, this chapter aims to provide a better understanding of the pitfalls and methodological contingencies of constructing indicators and the implications for researchers and practitioners. It proceeds as follows. The subsequent section employs the concepts of validity and reliability to discuss the methodological dilemmas and tensions inherent to the construction of governance indicators and how they apply to indicators of good governance in the context of sport. Next, the chapter employs these insights to analyse the strengths and weaknesses of the NSGO indicators.

This brief case study reveals important areas for further research that can aid the development of governance indicators in sport but also alerts practitioners to the inherent limitations of quantifying (good) governance in sport. The final section concludes these insights and explores the way forward.

The construction and use of indicators of good governance: A critical discussion

When engaging in a critical discussion of indicators of good governance, it is important to clear the conceptual cloud about indicators, composite indicators, indices, and benchmarking. An indicator can be defined as ‘a named collection of rank-ordered data that purports to represent the past or projected performance of different units’ (Davis, Kingsbury & Merry 2012, p. 5). Some common examples include a country’s employment rate, a company’s net sales, and an organisation’s level of compliance with environmental standards. When quantified, each indicator is represented by a value. The values of these individual indicators can subsequently be aggregated and weighted into a ‘composite indicator’, which is ‘a mathematical combination of individual indicators that represent different dimensions of a concept whose description is the objective of the analysis’ (De Lombaerde et al. 2011, p. 333). Composite indicators thus summarise complex and multi-dimensional realities by combining subindicators into a single number or index. An index, in turn, aids comparison of performance (‘benchmarking’) across units in one or more measured dimensions of governance (Nardo et al. 2005; Saisana & Tarantola 2002; Saisana, Tarantola & Saltelli 2005). Recent years have seen a proliferation of indicators and indices that measure the governance quality of both countries and different types of organisations. But their construction and use have not been uncontested.

Scholarly criticism of governance indicators can generally be grouped into two categories. A first group of scholars rejects the positivist ontological and epistemological assumptions of an objective and measurable social reality upon which the construction and use of indicators ultimately rests. Rather than focusing on the development of indicators, this group scrutinises the inherently political processes through which indicators are developed and used (Astleithner & Hamedinger 2003; see also the chapter by Claringbould et al. in this volume). A second group of scholars acknowledges the inherent limitations of measuring abstract dimensions of (good) governance, but nonetheless holds that assessing governance dimensions by measuring “imperfect proxies” is “both feasible and informative” (Kaufmann & Kraay 2008, p. 8). This group therefore focuses on developing methodologies that turn governance indicators into more accurate proxies for the abstract concepts they seek to measure (Gómez-Limón & Sanchez-Fernandez 2010; Marlier & Atkinson 2010; Nardo et al. 2005). To be clear, this chapter follows the reasoning of the second group only. While the reasoning of the first group has merit in that governance and its abstract dimensions are indeed generally unmeasurable, measuring proxies can still be meaningful, as

long as governance indicators' limitations are kept in mind. The remainder of this chapter therefore focuses on the methodological dilemmas and tensions that follow from the unavoidable subjectivity inherent to the construction of governance indicators. While it does not dwell on the political processes that underly the development of governance indicators, this issue is briefly addressed in the concluding discussion.

An increasingly large body of scholarly work scrutinises the construction of particular governance indices and explores the implications of particular methodological choices (Arndt & Oman 2006; Malito, Umbach & Bhuta 2018; Thomas 2010). Concerns about the technical dimensions of governance indicators and indices in this literature centre mostly on issues relating to validity and reliability. Though a comprehensive review of the existing tools is beyond the scope of this chapter, the following demonstrates how more general methodological concerns about governance indicators are also prevalent with regard to indicators of good governance in sport.

Validity

The first key criticism of governance indicators is their questionable validity. The central concern is construct validity, which focuses on whether indicators and indices measure what they intend to measure (Thomas 2010). Governance is a complex concept for which an agreed definition is lacking. This conceptual ambiguity has resulted in various operationalisations which, taken together, incorporate a diverse mix of macro, meso, micro, cultural, structural, input, process, or output factors. Consequently, different governance indices rely on varying criteria to assess the governance quality of countries and different types of organisations (Gisselquist 2012; Van den Berghe & Levrau 2003).

This diversity is not problematic *per se*. What is problematic, is that the construction of governance indices rarely rests on a basis of clearly specified concepts that are derived from a sound theoretical framework (Gisselquist 2012; Malito, Umbach & Bhuta 2018; Thomas 2010). This leads to overlaps and contradictions between dimensions and/or indicators, ends-means confusion (i.e. when input and process variables are inappropriately mixed with output or outcome variables), and ideological biases in governance indices (Malito, Umbach & Bhuta 2018; Nelken 2015). Most importantly, it implies that the producers of these indices have no basis for arguing that implementing the principles they put forward would impact some related variable in any way (Gisselquist 2014; Thomas 2010). When it is unclear *what* they intend to measure and how this correlates with other related variables, indicators do not constitute meaningful measurements of abstract dimensions of good governance (Thomas 2010).

The recently introduced sets of good governance indicators in sport appear to suffer from similar ills. Chappelet and Mrkonjic (2019) observe significant differences between these sets in terms of conceptual scope and operationalisation. They lament that most authors “do not explain why they chose certain principles

and not others, where these principles came from, and why they are set out in the way they are” (Chappelet & Mrkonjic 2019, p. 11). It thus remains unclear what aspects of governance these indicators seek to measure. Even when governance dimensions are defined, the theoretical framework that underpins these definitions is often implicit at best. Zintz and Gérard (2019, p. 56), for instance, simply note that the SIGGS governance dimensions and indicators were selected “after carefully analysing the literature”. In other words, while public discourse is characterised by the axiomatic use of good governance as a cure for the multiple failures of governance in sport, the expected impact of implementing different sets of principles remains ambiguous. This has two particular implications. First, it breeds confusion about the meaning of index scores. Pielke et al. (2020), for instance, question whether the Sports Governance Observer 2015 indicators accurately measure governance. The reason is their different, namely *output-oriented* understanding of (good) governance. Meanwhile, the Sports Governance Observer 2015 is an instrument designed to assess good governance in international sport federations based on an *institutional conceptualisation* of (good) governance (see also the case study below). Second, when the supposed effect of implementing particular principles is unclear, the appropriateness of these principles and the indicators that measure their implementation is questionable. Principles that are regarded as legitimate by some might actually have negative effects on the governance (outcomes) of sport organisations.

Reliability

The second major critique of existing governance indicators pertains to their reliability, that is, the extent to which they produce consistent results when applied by different people and across different case settings (Sarfaty 2015). Two particular reliability issues that preclude meaningful comparison emerge from the literature on governance indicators. The first issue pertains to the questionable reliability of the data that is systematically used to calculate indices that evaluate country governance. These data are typically either drawn from national databases or based on perceptions. The quality of the databases varies, and perceptions are typically influenced by external factors or changes and variations among evaluators (Malito, Umbach & Bhuta 2018). By contrast, corporate governance indicators often assess governance quality by triangulating publicly available data on firms with information gathered through interviews with executives (Van den Berghe & Levrau 2003). This produces more reliable data. Corporate governance indices that rely on companies’ self-reporting, however, suffer from self-overestimation and concealment (Sarfaty 2015). The second issue that negatively impacts the reliability of governance indicators, in both the public and private sector, is that they often do not rely on well-defined assessment criteria (Booyesen 2002; Sarfaty 2015). It goes without saying that indicators cannot be expected to produce consistent data, when there are no uniform measurement guidelines in place that can ensure their reliability (Sarfaty 2015, p. 111).

Similar reliability problems also appear prevalent with regard to governance indicators in sport. These indicators are assessed on the basis of data gathered via either self-evaluation, coding/analysis of publicly available information, or interviews or a combination of these (Chappelet & Mrkonjic 2019; Geeraert 2015, 2018b; Zintz & Gérard 2019). Self-evaluation is a very useful tool for stimulating reflection and learning. Yet it arguably produces the least reliable governance data of these three types (Geeraert 2019). When discussing the self-evaluation process that underpins the assessment of the SIGGS indicators, Zintz and Gérard (2019, p. 58) indeed acknowledge that “some organisations may have overestimated their level of good governance”. In addition, governance indicators in sport rarely specify precise criteria for measurement (Chappelet & Mrkonjic 2019). This poses problems when quantifying or categorising data. For example, ASOIF recognises that, when assessing the governance of the summer Olympic international federations on the basis of the KGP, the scoring criteria “lacked clarity in places” (ASOIF 2017, p. 7) such that “in many cases there was room for debate” (ibid, p. 8).

Functionality and transparency

While there is growing awareness of governance assessments’ validity and reliability problems, there is no commonly accepted method for attenuating these drawbacks (Cherchye et al. 2006; Gómez-Limón & Sanchez-Fernandez 2010; Marlier & Atkinson 2010; Nardo et al. 2005). A consensus has nonetheless emerged on a number of key methodological steps for constructing governance indicators, including (1) constructing a theoretical and conceptual framework; (2) selecting variables and constructing indicators; (2) devising a data-gathering strategy; (3) weighting and aggregating indicator scores; (4) robustness testing; and (5) devising a presentation strategy (Kaufmann, Kraay & Matruzzi 2011; Nardo et al. 2005). Yet, even the most ardent proponents of governance indicators stress that no measurement of good governance achieves perfect reliability and validity and that the construction of indicators inevitably involves a certain degree of inherent (methodological) subjectivity (Cherchye et al. 2006).

The literature therefore underscores the importance of functionality and transparency as guiding principles for dealing with subjective, methodological decisions. Functionality is, in essence, the practicality and usability of an indicator (De Peuter et al. 2007). In other words, when confronted with inescapable subjectivity, the researcher should aim to make indicators fit for the intended purpose (Cherchye et al. 2006; Nardo et al. 2005). Consequently, trade-offs have to be made between validity and reliability on one hand, and functionality on the other hand. For instance, while increasing the number of indicators might positively affect validity, measurement might become disproportionately costlier and even practically unfeasible. Validity will also suffer when indicators are formulated in broad terms, but this may be necessary in order to make them applicable to a diverse population of organisations (Astleithner & Hamedinger 2003). With regard to reliability, well-defined assessment criteria are necessary to

achieve consistent results. When indicators of good governance are formulated in an overly strict fashion, however, their application risks becoming too formalistic (Wymeersch 2006). Some flexibility is warranted, as (sport) organisations should have a degree of leeway to implement principles of good governance in accordance with their organisational culture and environmental context. Furthermore, consistent results may be essential for meaningful comparison, but arguably less important when indicators serve an educational purpose. For instance, broadly formulated indicators can be a very useful tool for stimulating debate within the context of an informal self-assessment. Indicators that are formulated in an overly strict fashion could be less fit for this purpose.

While there may be valid practical reasons for sacrificing some degree of reliability and validity, transparent reporting about the choices made in each of the key methodological steps is essential. It informs scholars as well as practitioners about the implications of (subjective) choices for the validity and reliability of the instrument (Gómez-Limón & Sanchez-Fernandez 2010, p. 1063; Marlier & Atkinson 2010, p. 289; Nardo et al. 2005, p. 22). In the absence of transparency about data and methodology, it is indeed impossible to assess what index scores actually mean. Given the governance effects of index scores, this lack of understanding is highly problematical.

Case study: The National Sports Governance Observer

To provide a better understanding of the pitfalls and methodological contingencies of constructing governance indicators in sport, this section critically explores the methodological choices that underpin the construction of the NSGO indicators (Geeraert 2018c). This comprehensive set of 274 indicators of good governance was the product of a recent project supported by the European Union Erasmus+ programme and managed by Danish NGO Play the Game. It measured and compared good governance in sport federations across nine European countries (Geeraert 2018c). Aiming to achieve methodological transparency, the section reports on the choices and trade-offs made in three key methodological steps. These are (1) constructing a theoretical and conceptual framework; (2) selecting variables and constructing indicators; and (3) weighting and aggregating indicator scores. The section ends with a critical reflection on the NSGO indicators, specifically in terms of validity, reliability, and functionality.

Theoretical and conceptual framework

The construction of the NSGO indicators followed the three-step approach for measuring abstract constructs as outlined by Thomas (2010). The first step requires defining governance, good governance, and its sub-components (Nardo et al. 2005). Because the NSGO was developed within the framework of a project that sought to improve sport organisations' internal structures, governance is defined from a narrow, institutional perspective (Treib, Bähr & Falkner 2007).

Institutions are understood here as sets of interconnected rules and practices that prescribe behaviour (Keohane 1988). In other words, governance is conceptualised as “a system of rules that shape the actions of social actors” (Treib, Bähr & Falkner 2007, p. 4).

Measuring governance, then, implies setting some normative benchmark against which the institutional design of organisations can be evaluated. In other words, it necessitates a clear conception of what constitutes ‘good’ governance in terms of appropriate rules and practices. The NSGO defines good governance as the embedding of transparency, democracy, internal accountability and control, and societal responsibility in the system of rules and practices that governs sport federations. Following Gisselquist, these four dimensions were chosen because they are recognisable to both academic and lay audiences, they are interrelated yet conceptually distinct, and they have theoretical utility since they build upon rich literatures (Gisselquist 2014, p. 517–18).

The second step outlined by Thomas (2010) involves formulating operational definitions, which allows translating the four abstract dimensions in observable institutional components (Gisselquist 2014, p. 519; Thomas, 2010). These definitions are elaborated in greater detail in [Chapter 2](#) in this volume. Transparency refers to institutional arrangements that make information available about an actor “allowing other actors to monitor the workings or performance of this actor” (Meijer 2014, p. 511). Democracy here entails a system of rules that establishes competition (electoral competition between political alternatives), participation (affected actors’ influence over collective decisions) and deliberation (fair and open debates aimed at mutual understanding) (Coppedge et al. 2011). For the third dimension, a narrow and internal perspective on accountability is adopted which entails a clear separation of executive, judicial, and supervisory powers on one hand, and an internal system that enables independent entities to monitor decision makers’ compliance on the other hand. Fourth and drawing on the Corporate Social Responsibility (CSR) definition by Van Aaken et al. (2013), societal responsibility is defined as the institutionalisation of pro-social behaviour. The notion of pro-social behaviour is employed here to capture those activities where sport federations employ their private authority to address societal issues related to sport.

Thomas’ (2010) third and final step involves making causal predictions about the implementation of the four components of good governance. Following a set of rational assumptions that are elaborated in detail in [Chapter 2](#) in this volume, it can be hypothesised that the joint implementation of the four components of good governance contributes to effective and impartial decision-making that is free from abuses of entrusted power for personal gain, to stakeholder trust, and to effective deployment of private authority to address societal issues.

Selecting variables and constructing indicators

Operationalising the four abstract dimensions of good governance requires replacing them with “intermediate objectives whose achievement can be observed and

measured” (Nardo et al. 2005, p. 5). Following a deductive approach, the NSGO indicators were constructed on the basis of the institutional rules and practices that best fit or express the four dimensions of good governance (Nelken 2015, p. 318). A review of more than 40 sets of good governance principles and recommendations issued by national governments, international organisations, and sport organisations as well as by the non-profit, corporate and cultural sectors was carried out. This led to the selection of 46 general practices, rules and procedures that fit the definitions of the four dimensions of good governance.

In order to make these 46 broad principles measurable, (sub-)indicators were constructed. For the sake of reliability, dichotomous variables were constructed, which have only two categories or levels (i.e. yes or no). These variables were quantified by giving the ‘yes’ category a value of 1 and the ‘no’ category a value of 0. This process resulted in 274 dichotomous indicators of good governance that evaluate the implementation of 46 broad principles of good governance, dispersed over the four dimensions. To further enhance reliability of the indicators, meta-data sheets were developed which explicitly detail the scoring criteria and possible data sources.

A three-fold strategy was followed in order to enhance the validity of the indicators. First, in line with the relevant literature, expert advice and feedback was sought and incorporated (De Peuter et al. 2007; Nardo et al. 2005). Both the selection of the 46 principles and the 274 indicators were informed by focus group discussions with Flemish sport federations and discussions with the academics and sport federation representatives that were part of the NSGO project. Second, the 274 indicators are both rules-based and outcome-based (Kaufmann & Kraay 2008). Though rules-based indicators are more straightforward to quantify, outcome indicators are able to capture bureaucratic practices that are not formalised but still relatively stable (Keohane 1988). Finally, a pilot study involving a mixed sample of small and large federations in nine European countries and Brazil ensured that the indicators apply to a broad range of different circumstances in a clear and unambiguous manner.

Weighting and aggregating indicator scores

In order to present the data in a meaningful, i.e. comparative way, the individual indicator values must be aggregated into one or more ‘composite indicators’. In the case of the NSGO, an ‘NSGO index’ score is calculated as follows. First, a score is calculated for each of the 46 principles on the basis of the average score of the underlying indicators. Second, each dimension (i.e. transparency, democratic processes, internal accountability and control, and societal responsibility) is assigned a score on the basis of the average scores of the principles that underlie the dimension. Finally, an NSGO index score is calculated on the basis of the average score of the four dimensions. Thus, like most composite indicators, the NSGO index, dimension scores, and principles score rely on equal weighting. This choice was informed by two considerations. First, the NSGO

index scores must be understandable for practitioners and equal weighting is the simplest weighting option. Second, weighting is especially subjective where academic research is still in its infancy (Nardo et al. 2005, p. 31). Equal weighting, while inherently subjective in its own right, is in this context arguably the least subjective solution.

While aggregation may aid comparison across federations, it is also important to ensure an objective comparison. Given that sport federations vary in terms of their administrative capacity and the governance-related risks they are subject to, the same governance standard cannot apply to all federations (Nardo et al. 2005). Three categories of sport federations were discerned on the basis of the samples selected by the research partners: small federations with fewer than ten full-time equivalent (FTE) employees, mid-sized federations (equal to or more than ten, but fewer than 30 FTE employees), and large federations (30 or more FTE employees). Accordingly, the indicators were divided into three categories, namely 'basic', 'intermediate', and 'advanced'. When indicators are not applicable, they are simply not considered in the calculation of scores.

Discussion: The limits of the National Sports Governance Observer

The key value of the NSGO is its ability to make a holistic and readily understandable diagnosis of potential institutional weaknesses and shortcomings of sport federations in four dimensions of good governance. A (modified) traffic light scoring system allows for an easily accessible yet nuanced interpretation of strengths and weaknesses. In addition, the Observer allows for collecting relatively reliable data that can be used for comparative research or for theory-testing purposes.

The instrument should be regarded as a barometer rather than a dogmatic tool that provides the ultimate assessment of good governance. Even though the indicators were constructed while paying specific attention to validity and reliability, the subjective choices that had to be made in the three steps outlined above are highly consequential. A brief, non-exhaustive overview of six particular limits to the instrument should suffice to inspire caution.

First, the instrument has a rational-structural bias. This is due to the underlying theoretical-conceptual framework's focus on material-institutional elements of governance. The NSGO index consequently provides an appreciation of the institutional qualities of sport federations in terms of four dimensions of good governance. This is the result of a trade-off with functionality: broadening the conceptual scope would multiply the already high number of indicators. However, this means that the NSGO index does not consider cultural-ideational (see also the chapters by Girginov and van Eekeren in this volume), personal (see also the chapter by Heres in this volume), or systemic (see also the chapter by van Bottenburg in this volume) factors. Even though the theoretical framework underpinning the indicators predicts a positive correlation between a high NSGO

index and organisational outcomes such as impact, effectiveness, or ethical conduct, the instrument does not provide a direct measurement of such outcomes.

Second, there is subjectivity involved in the selection of the four governance dimensions and the underlying principles and indicators. There is no point in denying that the selection of these dimensions is (implicitly) guided by normative considerations of “what ought to be” (Shapiro 2003, p. 2). These considerations are to some extent culturally inspired (see also the chapter by Girginov in this volume), as the four dimensions are in themselves considered “public values” in liberal democracies (see also the chapter by van Eekeren in this volume). Transparency, democracy, accountability, and societal responsibility may, however, have a different meaning in non-Western cultures.

Third, like most governance indices, the NSGO index relies on untested theoretical assumptions about institutional features (Fazekas, Tóth & King 2013). Because a theoretical framework accounts for the dynamics of implementing the four dimensions of good governance, achieving a high NSGO index is not meaningless. However, while these assumptions are always ‘true’ in the abstract world of theory, rigorous empirical testing is required to assess to what extent they hold in real-world circumstances and different (cultural) environments. The reality is, however, that the sport governance literature has thus far failed to address the impact of the implementation of principles of good governance on organisational outcomes (Parent & Hoye 2018).

Fourth, the choice for a total number of 274 indicators constitutes an imperfect compromise between practicality and validity. On one hand, the relatively high number of indicators arguably hurts the practicality and usability of the instrument. On the other hand, a case could equally be made for including more indicators in order to enhance the validity of the measurement. For instance, a sport federation that achieves a high score on the democracy dimension does not necessarily implement all relevant democratic processes (e.g. Coppedge et al. 2011).

Fifth, because there is no objective method for assigning weights, the weight of individual indicators and dimensions in the NSGO index is subject to endless dispute. In addition, the distinction between small, medium, and large sport federations in the construction of the NSGO index is equally subjective. This distinction, based on the number of FTE staff employed, is informed by expert opinion. As Parent and colleagues argue in their contribution to this book (see also the chapter by Parent et al. in this volume), this still constitutes a somewhat arbitrary benchmark since it is not based on empirical observation. It has, however, important implications, since ‘small’ federations are held to a lower standard with 66 of the 274 indicators deemed not applicable.

Sixth, because a degree of methodological subjectivity is inescapable, social processes likely play a significant role in the selection of indicators, dimensions, and weighting methods. There is no denying that this author’s own preconceived ideas about good governance, the project involvement of Play the Game as an advocacy organisation, the power relations between the project participants, and the participants’ attitudes and identities all have, to some extent, influenced the

construction of the instrument. Since these processes are largely unconscious, transparent reporting about them is nearly impossible, meaning that their impact remains indiscernible.

Conclusion

The recent proliferation of indicators of good governance in sport helps to fill important knowledge gaps, inform decision-making, and facilitate advocacy. At the same time, however, indicators that are poorly constructed, misinterpreted, manipulated or misused may induce bad governance. The brief case study of the NSGO indicators moreover demonstrates that even when particular attention is paid to validity and reliability, necessary trade-offs with functionality and inescapable subjectivity renders governance indicators inherently flawed.

For these reasons, it is indeed quite easy to criticise any set of governance indicators. Because abstract governance dimensions are unmeasurable and a fully objective and reliable measurement via proxies is impossible, one may even question the very existence of governance indicators. In my opinion, measuring governance can be meaningful and useful, but no measurement should be treated as final step. At the very least, scholars and practitioners should be attentive to several issues in order to mitigate the potential risks of using governance indicators.

Scholars have a special responsibility to ensure that risks are minimised and known. They should engage in the continuous improvement of existing indicators of good governance in sport, paying particular attention to know strategies for maximising validity and reliability. Two key areas for improvement regarding validity are constructing governance indicators on the basis of a sound theoretical framework and assessing the (unintended) impact of the implementation of governance principles across sport organisations in different (cultural) settings. There is room for diversity, as different indicator sets may have different functional goals, including education, advocacy, and inducing compliance. Different governance indices may therefore also quantify complementary aspects and dimensions of good governance in sport organisations. With regard to reliability, scholars should pay attention to developing clear and uniform measurement guidelines. In any case, they should report openly about contingent methodological choices and critically assess their implications. Yet, even the most transparent methodological reporting would fail to reveal the impact of the social and political processes that influence the construction and use of governance indicators. There is therefore a need for more critical academic research into these processes. For instance, on one hand, indicators are often used as a legitimisation strategy that privileges the status quo. On the other hand, they can be important tools for advocacy organisations that seek to delegitimise the status quo. Indicators can be manipulated to best serve either one of these aims.

Practitioners should handle governance indicators with care. They can be an important part, but they should not be the only component of any governance reform strategy. Governance indicators are particularly useful as first avenue to analysis of governance, as a way to suss out the situation. Subsequently, a more

in-depth investigation and discussion is required to establish the nature and severity of any established deficits. In any case, indicators should not provide a single rigid template for governance reforms. On one hand, the formalistic nature of indicators may induce box-ticking behaviours on the side of sport organisations rather than substantial reforms. On the other hand, these organisations should have a degree of leeway to implement principles of good governance in accordance with their organisational culture and environmental context. Awareness of the inherent shortcoming of governance indicators should also inspire a critical attitude towards the interpretation of governance indices. Practitioners should enquire about which aspects of governance are quantified by specific indicator sets. They should also inform themselves about the shortcomings of specific benchmarking instruments and demand transparency about contingent methodological choices and their implications.

References

- Arndt, C & Oman, C 2006, *Uses and Abuses of Governance Indicators*, OECD, Paris.
- ASOIF 2016, *ASOIF Governance Task Force (GTF) Report Approved by ASOIF General Assembly 2016*, ASOIF, Lausanne.
- ASOIF 2017, *First Review of IF Governance by ASOIF General Assembly 2017*, ASOIF, Lausanne.
- Astleithner, F & Hamedinger, A 2003, 'The analysis of sustainability indicators as socially constructed policy instruments: benefits and challenges of 'interactive research'', *Local Environment*, vol. 8, no. 6, pp. 627–40.
- Booyesen, F 2002, 'An overview and evaluation of composite indices of development', *Social Indicators Research*, vol. 59, no. 2, pp. 115–51.
- Chappelet, JL & Mrkonjic, M 2013, *Basic Indicators for Better Governance in International Sport (BIBGIS): An Assessment Tool for International Sport Governing Bodies*, IDHEAP, Lausanne.
- Chappelet, JL & Mrkonjic, M 2019, 'Assessing sport governance principles and indicators', in M Winand & C Anagostopoulos (eds), *Research Handbook on Sports Governance*, Edward Elgar, London, pp. 10–28. doi:10.4337/9781786434821.00008
- Cherchye, L, Moesen, W, Rogge, N & Van Puyenbroeck, T 2006, 'An introduction to "benefit of the doubt" composite indicators', *Social Indicators Research*, vol. 82, no. 1, pp. 111–45.
- Coppedge, M, Gerring, J, Altman, D, Bernhard, M, Fish, S, Hicken, A & Teorell, J 2011, 'Conceptualizing and measuring democracy: a new approach. *Perspectives on Politics*, vol. 9, no. 2, pp. 247–67. doi:10.1017/S1537592711000880.
- Davis, KE, Kingsbury, B & Merry, SE 2012, 'Indicators as a technology of global governance', *Law & Society Review*, vol. 46, no. 1, pp. 71–104.
- De Lombaerde, P, Dorrucchi, E, Genna, G & Mongelli, F 2011, 'Composite indexes and systems of indicators of regional integration', in P De Lombaerde, R Flores, L Lapadre & M Schulz (eds), *The Regional Integration Manual*, Routledge, Oxon, pp. 323–46.
- De Peuter, B, De Smedt, J, Van Dooren, W & Bouckaert, G 2007, *Handleiding beleidsevaluatie. Deel 2: Monitoring van beleid*, Steunpunt beleidsrelevant onderzoek bestuurlijke organisatie Vlaanderen, Leuven.

- Espeland, WN & Sauder, M 2007, 'Rankings and reactivity: how public measures recreate social worlds', *American Journal of Sociology*, vol. 113, no. 1, pp. 1–40.
- Fazekas, M, Tóth, IJ, & King, LP 2013, *Anatomy of grand corruption: a composite corruption risk index based on objective data*, CRC-WP/2013:02, The Corruption Research Center Budapest, Budapest. doi:[10.2139/ssrn.2331980](https://doi.org/10.2139/ssrn.2331980)
- Geeraert, A 2015, *Sports Governance Observer 2015. The Legitimacy Crisis in International Sports Governance*, Play the Game/Danish Institute for Sports Studies, Aarhus.
- Geeraert A 2018a, *Sports Governance Observer 2018. An Assessment of Good Governance in Five International Sports Federations*, Play the Game/Danish Institute for Sports Studies, Aarhus.
- Geeraert A 2018b, *National Sports Governance Observer, Final report*, Play the game/Danish Institute for Sports Studies, Aarhus.
- Geeraert, A 2018c, *National Sports Governance Observer. Final report*, Play the Game, Aarhus.
- Geeraert A 2019, 'The limits and opportunities of self-regulation: achieving international sport federations' compliance with good governance standards', *European Sport Management Quarterly*, vol. 19, no. 4, pp. 520–38. doi:[10.1080/16184742.2018.1549577](https://doi.org/10.1080/16184742.2018.1549577)
- Geeraert, A 2021, *National Anti-Doping Governance Observer. Final report*, Play the Game, Aarhus.
- Gisselquist, RM 2012, *Good Governance as a Concept, and Why This Matters for Development Policy*, United Nations University WIDER Working Paper.
- Gisselquist, RM 2014, 'Developing and evaluating governance indexes: 10 questions', *Policy Studies*, vol. 35, no. 5, pp. 513–31. doi:[10.1080/01442872.2014.946484](https://doi.org/10.1080/01442872.2014.946484)
- Gómez-Limón, JA & Sanchez-Fernandez, G 2010, 'Empirical evaluation of agricultural sustainability using composite indicators', *Ecological Economics*, vol. 69, no. 5, pp. 1062–75.
- IOC 2008, *Basic Universal Principles of Good Governance of the Olympic and Sports Movement*, IOC, Lausanne.
- Kaufmann, D & Kraay, A 2008, 'Governance indicators: where are we, where should we be going?', *The World Bank Research Observer*, vol. 23, no. 1, pp. 1–30.
- Kaufmann, D, Kraay, A & Matruzzi, M 2011, 'The worldwide governance indicators: methodology and analytical issues', *Hague Journal on the Rule of Law*, vol. 3, no. 2, pp. 220–46.
- Keohane, RO 1988, 'International institutions: two approaches', *International Studies Quarterly*, vol. 3, no. 4, pp. 379–96.
- Malito, DV, Umbach, G & Bhuta, B 2018, *Palgrave Handbook of Indicators in Global Governance*, Palgrave MacMillan, New York.
- Marlier, E & Atkinson, AB 2010, 'Indicators of poverty and social exclusion in a global context', *Journal of Policy Analysis and Management*, vol. 29, no. 2, pp. 285–304.
- Meijer, A 2014, 'Transparency', in M Bovens, RE Goodin & T Schillemans (eds), *The Oxford Handbook of Public Accountability*, University Press, Oxford, pp. 507–24.
- Merry, SE 2016, *The Seductions of Quantification: Measuring Human Rights, Gender Violence, and Sex Trafficking*, University of Chicago Press, Chicago.
- Nardo, M, Saisana, M, Saltelli, A & Tarantola, S 2005, *Tools for Composite Indicators Building*, European Commission, Brussels.
- Nelken, D 2015, 'Conclusion: contesting global indicators', in SE Merry, KE Davis & B Kingsbury (eds.), *The Quiet Power of Indicators. Measuring Governance, Corruption, and Rule of Law*, Cambridge University Press, Cambridge, pp. 317–37.

- Parent, MM & Hoye, R 2018, 'The impact of governance principles on sport organisations' governance practices and performance: a systematic review', *Cogent Social Sciences*, vol. 4, no. 1, pp. 1–24. doi:[10.1080/23311886.2018.1503578](https://doi.org/10.1080/23311886.2018.1503578)
- Pielke, R, Harris, S, Adler, J, Sutherland, S, Houser, R & McCabe, J 2020, 'An evaluation of good governance in US Olympic sport National Governing Bodies', *European Sport Management Quarterly*, vol. 20, no. 4, 480–99. doi:[10.1080/16184742.2019.1632913](https://doi.org/10.1080/16184742.2019.1632913)
- Saisana M & Tarantola S 2002, *State-of-the-art report on current methodologies and practices for composite indicator development*, Report 20408, European Commission, Brussels.
- Sarfaty, GA 2015, 'Measuring corporate accountability through global indicators', in SE Merry, KE Davis & B Kingsbury (eds.), *The Quiet Power of Indicators. Measuring Governance, Corruption, and Rule of Law*, Cambridge University Press, Cambridge, pp. 103–31.
- Saisana, M, Tarantola, S & Saltelli, A 2005, 'Uncertainty and sensitivity techniques as tools for the analysis and validation of composite indicators', *Journal of the Royal Statistical Society Series A*, vol. 168, no. 2, pp. 1–17.
- Saltelli, A 2007, 'Composite indicators between analysis and advocacy', *Social Indicators Research*, vol. 81, no. 1, pp. 65–77.
- Shapiro, I 2003, *The State of Democratic Theory*, Princeton University Press, Princeton.
- Thomas, MA 2010, 'What do the worldwide governance indicators measure?' *European Journal of Development Research*, vol. 22, no. 1, pp. 31–54.
- Treib, O, Bähr, H & Falkner, G 2007, 'Modes of governance: towards a conceptual clarification', *Journal of European Public Policy*, vol. 14, no. 1, pp. 1–20. doi:[10.1080/135017606061071406](https://doi.org/10.1080/135017606061071406)
- Van Aaken, D, Splitter, V & Seidl, D 2013, 'Why do corporate actors engage in pro-social behaviour? A Bourdieusian perspective on Corporate Social Responsibility', *Organization*, vol. 20, no. 3, pp. 349–71.
- Van den Berghe, LAA & Levrau, A 2003, 'Measuring the quality of corporate governance: in search of a tailor-made approach?', *Journal of General Management*, vol. 28, no. 3, pp. 71–86.
- Wymeersch, E 2006, 'The enforcement of corporate governance codes', *Journal of Corporate Law Studies*, vol. 6, no. 1, pp. 113–38.
- Zintz, T & Gérard, S 2019, 'SIGGS: a European project for national Olympic committees and national sport federations' in M Winand & C Anagnostopoulos (eds.), *Research Handbook on Sport Governance*, Edward Elgar, Cheltenham, pp. 53–71.