A Mixed Methods Approach to Compare Elite Sport Policies of Nations. A Critical Reflection on the Use of Composite Indicators in the SPLISS Study

Prof. Dr. Veerle De Bosscher

Reference:

De Bosscher, V. (2018). A Mixed Methods Approach to Compare Elite Sport Policies of Nations. A Critical Reflection on the Use of Composit Indicators in the SPLISS Study. Sport in Society: Cultures, Commerce, Media, Politics. Special issue: Research Methodologies and Sports Scholarship, editors James Skinner & Terry Engelberg, 21, 2, pp. 331-355, pp. 331-355. http://www.tandfonline.com/doi/pdf/10.1080/17430437.2016.1179729?needAccess=true

Institutional address: Vrije Universiteit Brussel

Faculty of Physical Education and Physiotherapy

Prof. Dr. Veerle De Bosscher

Pleinlaan 2, 1050 Brussels, Belgium

Corresponding author : Veerle.de.bosscher@vub.be

Abstract

This paper discusses the utility of mixed methods research in international comparative studies on elite sport policies, and (quantitative) composite indicators (CIs) in particular. It illustrates how complex and large amounts of data in 15 nations have been objectified into easily understood formats, CIs. Using a nine Pillar model, data were collected through a research inventory and surveys completed by 3,142 elite athletes, 1,376 coaches and 241 performance directors. 96 critical success factors and 750 sub-factors were aggregated into a CI.

The paper shows how CIs are helpful in identifying a possible (non)relationship between elite sport policies and success, in facilitating interpretation and comparison, and in understanding differences and convergences in elite sport systems. However, there are a number of drawbacks, for example understanding elite sport policies as part of a broader social, cultural and political context. Complementary qualitative analysis is necessary to interpret elite sport policies of nations.

Keywords: policy evaluation, mixed methods research, international comparison, elite sport policies, benchmarking, composite indicators, competitiveness, SPLISS

Introduction

Mixed methods have been embraced by a growing list of academic areas, including psychology, social work, nursery, medicine, health sciences, management, organisational studies, evaluation and education (Creswell 2009, Tashakkori and Teddlie 2003). Numerous studies have illustrated how mixed methods designs can be used to examine complex social and health issues (Vrkljan 2009). Despite the broad acceptance of the technique, mixed methods have not yet been fully explored as a means to reduce problems related to (in)comparability in international comparative policy research. Researchers have been mixing qualitative and quantitative approaches for decades in many subjects, including sport management, but to put both forms of data together as a distinct research design is relatively new (Creswell and Plano Clark 2007). This is particularly true in sport policy and sport management research, where mixed methods research is still rarely used, poorly legitimised and often weakly designed (Rudd and Johnson 2010, van der Roest, Spaaij, and van Bottenburg 2015).

The purpose of this paper is to demonstrate the utility of mixed methods research in international comparative studies and, notably, those concerned with elite sport policy. This paper will illustrate how a blend of mixed research designs was used to make an international comparative study in 15 sample nations on the sports policy factors leading to international sporting success (SPLISS). This paper will particularly focus on how, in addition to qualitative data, (quantitative) composite indicators (CI) are useful for comparing and objectifying large amounts of international data on elite sport policies into easily understood formats and for identifying possible relationships between elite sport policies and international sporting success. CIs are synthetic indices of individual indicators that have been increasingly used to rank countries in various performance and policy areas (Freudenberg 2003). This quantifying technique will allow analysts to discern and to show regularities or peculiarities in qualitative data they might not otherwise see (Sandelowski, Voils, and Knafl 2009). However, the paper

highlights that in an elite sport policy setting, these CIs are not an isolated measurement or ranking system by themselves; they are rather a supportive and tangible way of understanding elite sport policies more broadly in relation to sporting success. Social phenomena such as elite sport policies are complex and reality is more than what we merely measure with CIs. As the CIs must be interpreted in relation to the wider structural mechanisms that produce the observed effects, this study combines qualitative data and interpretations with quantitative scores (CIs). Accordingly, this paper will critically assess the methodological strengths and weaknesses as well as the challenges of quantitative measurements such as CIs for the evaluation of elite sport policies in order to illustrate the usefulness of a mixed methods design.

The methods explored in this paper are part of a large scale project, called SPLISS 2.0 (Sports Policy factors Leading to International Sporting Success). The aims were to improve the development of theory concerned with the key success factors in elite sport policy as well as the methods employed to compare elite sport policies more objectively and less descriptively (De Bosscher et al. 2015). In this project, a complex mixed methods design was constructed, with multiple stages and a combination of sequential and concurrent mixed methods phases. Mixed methods were integrated at the level of data collection, analysis, interpretation and reporting, using the design models proposed by Creswell and Plano Clark (2007). As Johnson and Onwuegbuzie (2004) note, the integration of quantitative and qualitative data helps to reduce drawbacks of both methods; it provides a better understanding of the problem than if either dataset had been used alone, and it also offers a more comprehensive picture by taking account of the trends and generalizations as well as in-depth knowledge of participants' perspectives.

International Comparative Research on Elite Sport Policies

International comparative research, particularly in high performance sport, is complicated, because sport is intertwined with commercial, political, social and cultural factors. According to Houlihan (2013), this explains why, despite a steady convergence of elite sport systems in a

globalized world, a high level of variation between nations persists. As such, there are many extraneous and uncontrollable factors that make comparability of elite sport policies, systems and management practices difficult. De Bosscher (2007) summarised six key points that need to be considered in international comparative research: (1) there is a need for a clear definition of concepts; (2) little availability of comparable data may lead to misconceptualisation and misinterpretation; (3) isolation of data from their broader (cultural and historical) context may lead to over-simplification, cluster of dissimilar phenomena and failure to recognise important differences; there is a need to explain 'why'; (4) danger of/ difficulty in identifying cause and effect; (5) implementation of results/policies may be poorly replicated to other contexts; (6) ignoring important differences and leads to clustering of unlike phenomena. These key points explain why there is a lack of standardisation of methods used for comparative research and, as such, there are no universal or perfect methods in cross-national studies (Henry et al. 2005) addressing all these issues. Therefore, typical problems in comparative research cannot be eliminated, but they need to be reduced as much as possible (De Bosscher et al. 2010).

As a consequence, most (elite) sport policy studies are qualitative and descriptive in nature. Table 1 provides an overview of the growing number of international comparative studies (book publications) over the past decade that have enhanced our knowledge on elite sport development in different nations and contributed to a better understanding of elite sport systems and the factors that shape policy. These studies are qualitative in nature and focus on analysing elite sport systems as a complete and interacting whole rather than an assembly of distinct and separate elements in order to avoid the risk that this might obscure understanding (Andersen, Houlihan, and Ronglan 2015). The two SPLISS studies in the table (in 2008 and 2015) adopted a mixed methods approach combining qualitative data collection and analysis with quantitative CIs to compare nine policy dimensions ('Pillars') and critical success factors (CSF) of nations, through the involvement of key stakeholders in elite sport in the policy evaluation, such as elite

athletes, elite coaches and performance directors. While all these studies had broadly analogous goals to analyse and compare elite sport systems of nations and identify similarities and divergence, 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). Mixed methods can illuminate issues that cannot be explained by qualitative or quantitative approaches alone and provide more comprehensive evidence for studying a problem (Creswell and Plano Clark 2007) and they are particularly useful to analyse relationships (Sandelowski, Voils, and Knafl 2009). Therefore, in addition to qualitative data analysis, SPLISS applied CIs, reflecting economic competitiveness studies. This will be explained more in depth in the next section.

INSERT TABLE 1 NEAR HERE

Composite Indicators

Composite indicators (CIs), which are synthetic indices of individual indicators (Freudenberg 2003), have been increasingly recognised as a useful tool in policy analysis and public communications and to compare country performance (Nardo et al. 2008). They are generally used to summarise a number of underlying individual indicators or variables. An indicator is a quantitative or qualitative measure derived from a series of observed facts that can reveal relative position in a given area and, when measured over time, can point out the direction of change (Freudenberg 2003). Using composites, countries have been compared with regard to their competitiveness, innovative abilities, degree of globalisation and environment sustainability (Freudenberg 2003). More than 190 studies using CI have been registered (Berger and Bristow 2009). Drawing on these mainstream economic studies, a number of researchers in sport management have started to use composite indicators to

compare nations. They share a focus on the relationship between the internal characteristics of an organisation and its performance (Truyens et al. 2015). Table 2 provides a selection of four examples of economic studies measuring competitiveness and four sport management/policy studies, including the SPLISS 2.0 study which applied principles of CIs to measure their goals in addition to qualitative data. This table illustrates that all these studies seem to have applied similar basic principles of the composition of CIs in terms of the design and the following procedures (Linssen 1998, Freudenberg 2003, Nardo et al. 2008): (1) Developing a theoretical framework and generating the determinants for the composite under a fitness-for-purpose principle; (2) Data selection: identifying and developing relevant variables in indicators and sub-indicators; the use of proxy variables should be considered when data are scarce; consideration and interpretation of missing data; (3) Scoring each indicator, standardising variables to allow comparisons and, in some cases, weight variables; (4) Aggregate the scores after normalisation according to the underlying theoretical framework; check for robustness; (5) Comparing the scores; presentation and visualisation; interpretation of the theoretical framework, identifying differences between nations.

However, while the general design of the studies in Table 2 is essentially similar, each of these indices uses different conceptualisation methods, scoring methods, standardisation methods and weightings. The differences confirm that there does not exist a blueprint or practical guideline to develop CIs in one particular way. For example, in the World Competitiveness Yearbook (Institute for Management Development (IMD) 2012), 61 economies are analysed and ranked on over 300 criteria that are grouped into 20 factors and then regrouped into four competitiveness determinants. A continuous scaling method is used, where data from a 1–6 scale are converted to a 0–10 scale and standard deviation values calculated to determine rankings. In the Global Competitiveness report (World Economic Forum (WEF) 2007), three component indices are calculated on the basis of 35 sub-indices and,

with a score class method, the data are converted to a scale of 1–7 (Onsel et al. 2008). In other indices, such as the Economic Freedom of the World index (Fraser Institute 2005), 21 components in five major areas are incorporated into the index, which is made up of several sub-components. Further, in the Heritage Foundation Index of Economic Freedom, 50 independent variables are divided into ten broad factors of economic freedom (Ochel & Röhen, 2006). Hard (available) data are collected and, in most cases, also soft data through surveys among business executives. These data are scored and then aggregated into a final score for each dimension, possibly after weighting.

The sport management CIs use similar constructs, but mainly differ from these economic indices in the smaller number of nations being compared. This is likely to be related to the complexity of a sport setting that is embedded in the aforementioned broader cultural context, where beliefs, norms and values have been shown to have had a marked impact on the character of sport management/policy (Houlihan & Green, 2008). Sport management studies also have a different approach to the conceptualisation of their work. Most of them use a triangulation of methods and their measurement is preceded by research on the identification of a framework and criteria. In contrast, the economic competitiveness studies mostly make use of hard data that are readily available(online), such as factors influencing unemployment rate, employment changes, GDP per capita and potential growth (Berger and Bristow 2009), or use a proxy when data are scarce (Nardo et al. 2008), allowing them to have access to a larger number of nations. The main critique of these studies is that they fall short in the conceptualisation, which is linked to the lack of a sound overall model of national competitiveness to guide this process (Berger, 2009). The sport management studies usually collaborate in their project with a local researcher or organisation to collect data, which is time consuming and intensive. In addition to hard data collection, most economic studies include survey data among business executives, which is in the sport management studies only present in the SPLISS methodology. Finally, the CIs in sport

management studies have a different purpose, as composing an index is not a purpose as such, but is used as an informative tool to evaluate performance or competitive advantage more broadly (Truyens et al. 2015).

In conclusion, the overview in Table 2 illustrates that there is no uniform method to compose a composite indicator, but many studies have used CIs for similar goals, in other words, to provide simple comparisons to illustrate complex and sometimes elusive issues. The next section will explore the methods used in the SPLISS 2.0 study in more depth in order to illustrate how CIs (quantitative scores), which were composed of both qualitative and quantitative data, enriched the qualitative and descriptive policy evaluation in 15 nations.

INSERT TABLE 2 NEAR HERE

Methods SPLISS 2.0

The four major types of mixed methods designs, as defined by Creswell and Plano Clark (2007), seem helpful in addressing multifaceted research problems and assisting researchers with the design selection process. However, the challenge remains when a research problem cannot be addressed using one specific design type (Vrkljan 2009). Therefore, the SPLISS project that started in 2002 blends features from various design types. As illustrated in Figure 1, the projected used a three phase sequential exploratory design (Creswell and Plano Clark 2007). In the first phase (developing the SPLISS model), the design was exploratory because of the premise that "an initial qualitative exploration is needed for the reason that measures or instruments are not available, there is initially no guiding framework or theory and the variables are unknown" (Creswell and Plano Clark 2007). During this (mainly) qualitative phase, a nine Pillar conceptual model was developed because of the lack of an existing theory on the sports

policy factors that influence international sporting success (De Bosscher et al. 2006). Each pillar consisted of Critical Success factors (144 in total). The second phase, reflected in the name SPLISS 1.0, built on the results obtained in the first phase and was therefore sequential (Creswell and Plano Clark 2007). A concurrent triangulation design was embedded consisting of obtaining concurrent qualitative and quantitative data on the elite sport policies of six sample nations in order to test the conceptual framework in an empirical environment. Subsequently, the data were qualitatively analysed, compared and reported. Different from other elite sport policy studies, the study mirrored economic competitiveness studies that transform qualitative and quantitative data into a quantitative scoring system (using CIs) to assist the qualitative descriptive analysis.

The third phase (SPLISS 2.0), developed between 2010 and 2015, is a follow-up study comparing 15 nations. We take this project beyond SPLISS 1.0 by collecting more information about certain Pillars, developing a more comprehensive scoring methodology and obtaining deeper insights into the relationship between elite sport policies and sporting success of nations. This further contributed to the validation of the theoretical model. A number of CSFs were merged; a total number of 96 CSFs have been evaluated in the nine pillars. The project was non-funded and has been realised by the collaboration of the local researchers who took care of their own data collection, using the SPLISS instruments and methods.

The processes involved with these three phases will be further explained in the next section. We also refer to previous publications for more detailed information about the mixed methods used in SPLISS 1.0 (De Bosscher et al. 2006, De Bosscher et al. 2010) and the validity procedures (De Bosscher 2016).

INSERT FIGURE 1 NEAR HERE

---

1. The SPLISS model and CSFs

SPLISS (De Bosscher et al. 2006) clusters all factors within sport policy that can contribute to success (outputs) in nine policy dimensions, called 'Pillars', situated at two levels: inputs (Pillar 1) and throughputs (Pillars 2–9). Inputs are reflected in Pillar 1 as the financial support for sport and elite sport. Countries that invest more in (elite) sport can create more opportunities for athletes to train under ideal circumstances to develop their talent. Throughputs are the policy actions that script and deliver the processes in elite sport policies ("what" is invested and "how" it is used) that may lead to increasing success in international sport competitions. They refer to the efficiency of sport policies; that is, the optimum way the inputs can be managed to produce the required outputs. All of the Pillars 2–9 are indicators of the throughput stage.

Each Pillar is operationalised into measurable sub-dimensions. These are the critical success factors (CSFs) that identify *what* characterises successful elite sport policies, and also *how* these different dimensions (Pillars) can be developed. CSFs within each Pillar are elements that can drive the Pillar forward. A total of 96 CSFs and nine Pillars were measured in SPLISS 2.0, as shown in Table 3. We refer to previous publications for an overview of the model and its CSFs (De Bosscher, De Knop, and van Bottenburg 2009, De Bosscher et al. 2015).

INSERT TABLE 3 NEAR HERE

The concurrent qualitative and quantitative data collection and analysis have already briefly been exemplified in Figure 1 and Table 2 in comparison to other CI studies, and will be further explained in the next section.

2. Sample nations

When the SPLISS 2.0 project was announced, any interested nation was invited to participate under the condition that they were able to collect the comprehensive data set and follow the research protocol. Table 4 shows the 15 sample nations, with their population and wealth (expressed as GDP per capita), as these factors explain over 50% of international sporting success (De Bosscher et al. 2015, De Bosscher et al. 2006). These sample nations represent 16 sport systems, from 13 nations and three regions[1]: Flanders and Wallonia (representing Belgium) and Northern Ireland, one of the four countries of the United Kingdom (UK). As the UK did not take part in the study, Northern Ireland was therefore seen as a 'nation in its own right' within the project. Japan and Brazil are the largest countries of the sample, with populations of 127 and 203 million inhabitants respectively. Estonia and Northern Ireland have the smallest populations. In terms of wealth, Switzerland exceeds all nations, followed by the three medium sized populated nations (Canada, Australia and the Netherlands). Brazil's GDP per capita is much lower than that of the other nations.

INSERT TABLE 4 NEAR HERE

---

3. Protocol

The national SPLISS research partner collected the data locally in each country, using pre-defined research instruments. A total of 58 researchers and 33 policy makers collaborated in this project, with one coordinator per nation. Taking the complexity of international comparative research into consideration, the study was coordinated by the lead researcher from Belgium in collaboration with an international consortium group from three countries (Australia, the Netherlands and United Kingdom). Comparability of data and the reliability of the comparison was a major concern of the research consortium. Researchers received a research protocol that provided guidance on the process of data collection, aiming to

standardise data gathering procedures. All documents were provided through a joint web platform. Several international meetings were organised to fine-tune the data collection and identify possible gaps in the research methodology. Various phases of interpretative validation were included during the process, when the local researchers were asked to check the scores and policy description in regard to data interpretation by the authors.

4. Qualitative and quantitative data collection

Compared with economic studies, only limited comparable data on elite sport policies are available for nations. The Pillars and the 96 CSFs were operationalised through two types of research instruments used to collect complementary data, as outlined below. An overview of the methods is also shown in Table 2.

(1) An overall elite sport policy inventory, which was a comprehensive research instrument in its own right, was used to collect mainly qualitative data on all Pillars identified. It was completed by the relevant researchers in each country through interviews with policy agencies and analysis of existing secondary sources, such as policy documents. For each Pillar, the 96 CSFs were divided into open-ended and closed questions. The open-ended questions primarily sought to gain an insight into each country's policy system for each pillar and the presence of resources and how these resources were used. Closed questions (yes/no) were added to ensure a degree of comparability for the various sub-criteria. In order to enable subsequent quantification, questions were subdivided in several dichotomous sub-questions for each of which space was left for additional comments and every pillar ended with two specific questions concerning: 1) the main strengths and weaknesses of the findings from each pillar; and 2) the researchers' suggestions for improving the elite sport policies in their country. These open sections allowed the researchers to provide further details about possible additional criteria, which had not been included in the inventories, but which were issues specific to their

country. The inventories contained 184 pages with 212 open ended and closed questions covering all the nine Pillars (De Bosscher et al. 2015).

(2) The elite sport climate survey, completed by 3142 athletes[2], 1376 coaches and 243 performance directors (national governing bodies) of each nation, served two purposes: (1) to gather (mainly quantitative) information on indicators or "facts" that cannot easily be measured (using dichotomous questions) (De Pelsmacker and Van Kenhove 1999); and (2) to measure success indicators as they are perceived by their primary users (using a five point Likert scale), referring to the marketing services literature and the effectiveness literature which states that the primary stakeholders in sport organisations should be involved (Chelladurai 2001, Shilbury and Moore 2006). For example, while "effective communication" is a critical success factor for Pillar two (organisation and structure), it is not easy to evaluate and quantify, subject to assessment by the primary users.

The fact that the project was highly dependent on the cooperation of sports authorities and Olympic Committees, which had not necessarily endorsed the research in all countries, made it challenging to access all three respondent groupings in some countries. In some countries, it was hard to collect all information for all Pillars. Estonia only completed the Pillar 1 inventory and South Korea did not complete Pillars 3 (participation), 4 (talent), 7 (coaches) and 8 (international competition). Furthermore, France was unable to participate in the surveys due to final approvals arriving after the data collection deadlines. Obviously, prudence is needed in the comparison and, wherever information is incomplete, this will be shown.

5. Data analysis

5.1 qualitative data

In the inventories, data collected from document analysis and interviews were recorded and subsequently processed and interpreted by the local researcher. Reading and interpreting these

inventories was important to gain a deeper understanding of each sport system and the development of each CSF. These data all together contained over 3000 pages of mainly qualitative information, which was coded and grouped by CSF and compared across the 15 nations.

5.2 Transformation into quantitative data to develop a scoring system (CI)

The use of CIs is the central subject in this paper. Reflecting economic competitiveness studies, a total of 750 sub-factors (from quantitative and qualitative data from the inventories and the surveys), were allocated a score between 0 and 1, and aggregated into the 96 CSFs and subsequently into a CI for each of the nine Pillars. Depending on the source (elite sport climate survey or sport policy inventory) and type of question (open ended, dichotomous or assessment), the standards for this scale differed, as explained below.

The most complex ratings were derived from the overall sport policy inventory, because qualitative information on the elite sport systems for each Pillar had to be transformed into a score. These (mostly) open ended questions were grouped and assessed in terms of availability of the criterion in a stronger or weaker form, to indicate the level of development. As shown in Table 2, a score class method was used, whereby qualitative indicators that represent the development of specific policy characteristics and sub-characteristic are scored by a dummy value (0 or 1). The more sub-characteristics that could be identified for a specific resource, the higher the value. For each CSF, all standards and ratings were discussed within the consortium group until consensus was reached. For quantitative data from the overall sport policy inventory (e.g. elite sport expenditures), data were standardised. "Z-scores" were created for all quantitative data sets, allowing fair comparisons between different types of data. Each data point was given a score based on its distance from the mean average of the entire data set, where the scale is the standard deviation of the data set. Subsequently, the Z-score was turned into a "cumulative probability score" to arrive at the final totals (between 0 and 1) for each CSF.

More details on the composition of CIs can be found in previous publications (De Bosscher et al. 2010, Truyens et al. 2015). Relevant examples are provided in Table 2.

In the elite sport climate survey, quantitative data were available based on two types of question: dichotomous questions (yes/no) and ratings on a five point Likert scale (ordinal). For the dichotomous questions, absolute standards were used to calculate the scores (the percentage of 'yes' answers divided by 100). For the 1–5 Likert scale (perceived) questions, ratings were calculated by multiplying the response values respectively by 1 (highly developed), 0.75 (sufficiently developed), 0.5 (reasonably developed), 0.25 (insufficiently developed) and 0 (not developed). This resulted in a score that lies between 0 and 1.

The sub-factor scores were totalled for each CSF and then aggregated into a total percentage score for each Pillar. The total score was allocated a conditional formatting, ranging from a low level of development (dark grey) to a high level of development (light grey).

Finally, some criteria were weighted to reflect the consortium's view of their relative importance. These weightings were needed primarily because not each CSF was measured by the same number of questions, and to "lock in" the impact of each CSF on the overall score.

Results

In the context of policy analysis at national and international levels, indicators are useful for identifying trends in performance and policies and for drawing attention to particular issues (Freudenberg 2003). The first section of the results will illustrate the CI-scores of the 15 nations and the interpretation of the scores for informing theories and policies, particularly with regard to the relationship between policies and success. The second section illuminates some elements that do not emerge from the scores and that can only be explored by additional qualitative analysis.

Comparing National Elite Sport Policies - What CIs CAN do

Figure 2 displays the CI scores for all 15 nations in the nine Pillars. As a reminder, these are the aggregated scores within each Pillar (96 CSFs, 750 sub-factors) of data collected through the inventories (by the researchers – over 3000 pages) and the elite sport climate surveys (by the athletes, coaches and performance directors). As SPLISS aims to explore the relationship between elite sport policy systems (inputs and throughputs) and success in international competitions (outputs), the countries are ranked according to their success in summer sports, measured as market share of medals[3] in the 1,065 Summer Olympic Games and World Championships events contested between 2009 and 2012. We need to take note of the fact that some countries heavily focus on winter sports, notably Canada, Switzerland and Finland. Therefore the winter sport market share is shown in the second column. Countries with incomplete datasets are indicated by the * and the scores are presented between brackets.

As a general overview of Figure 2, it can be noted that higher performing countries in summer sports tend to have higher scores on the nine Pillars. There are some exceptions, such as Brazil, scoring low on most Pillars (except Pillars 1 and 8); and across all countries, low scores on Pillar 3 (sports participation) and Pillar 4 (talent identification and development). In the bottom half of Figure 2, some less successful countries still display higher scores in Pillar 4 (talent identification and development) and Pillar 5 (athletic career and post career support). The main point of note is that the overall scores and (grey scale) traffic lights may increase insights into the relationship between policies and success. A correlation matrix is shown in Table 5. Overall, most Pillars correlate positively and significantly with success, either in summer or winter sports: Pillar 1 (financial support), Pillar 2 (structure and organisation), Pillar 7 (coaches) and Pillar 9 (scientific research) are the four Pillars that correlate significantly with sporting success for both summer and winter sports. Pillar 6 (facilities) and Pillar 8 (international competition) correlate significantly with summer sports only. These correlations do not indicate any causality but show that the variation in the two data sets is similar. For example, Pillar 1

(financial support) seems to be the best predictor of outputs (success), but it is also possible that elite sport success causes financial support. Meanwhile, the Pillar scores reveal that financial support does not guarantee success, as is evidenced by the scores in the Netherlands, that does well in most Pillars but with a relatively small budget (Pillar 1); Brazil, at the other end of the scale, has high elite sport expenditure and low scores on most Pillars. For winter sports, the relationship between success and the nine Pillars is less pronounced. This may be attributable to winter sports being more specialised than summer sports and that fewer nations prioritise winter sports. A further external factor that cannot be influenced by policy makers is the dependency of winter sports on the natural (landscape and climate) environment, such as mountains and snow.

Since the intention of this article was to illustrate the use of CIs to compare elite sport policies, we cannot dwell on the detailed explanation for each Pillar and comparison of nations in relation to success. Figure 2 is a one page summary of more than 3000 pages of inventory data and survey results of 3142 elite athletes, 1376 elite coaches and 241 performance directors that completed the elite sport climate survey. As such, the Figure exemplifies the ability of CIs to integrate large amounts of information into an easily understood format for a general audience and thus enable users to compare complex dimensions effectively.

INSERT FIGURE 2 NEAR HERE

INSERT TABLE 5 NEAR HERE

These results also exemplify that CIs have other advantages of helping to interpret the data; for instance, they can produce statistical insights into the relationship between policies and success, which in turn contributes to establishing the criterion validity of the SPLISS model.

When we delve deeper into the 96 CSFs that are the building blocks of the nine Pillars, we find that 22 factors correlate significantly with success either in summer or in winter sports (at the 0.05 level); 50 have a correlation higher than 0.3.

Furthermore, deconstructing CIs into the scores of different CSFs can show the contribution of subcomponents and individual indictors. Although it is not shown in Figure 2, it can reveal what is driving the CI results and if nations are dominating certain factors, as those countries may disclose different strengths and weaknesses in different CSFs within the same Pillar. Moreover, scores and CSFs can be linked to other (measurable) phenomena, like population or wealth. For example, the results brought to light that higher scores were notable in smaller countries (both in population and area) in Pillar 4 (talent identification and development), and that grassroots and elite sport expenditure (government, lotteries and NOCs) does not relate significantly to the wealth of the countries in the SPLISS sample. Finally, the scores provide insights into the inter-relationships between Pillars. For example, an interesting point of note was that precisely those countries that were identified as being the most efficient in Pillar 1 (Australia, Japan, France and the Netherlands for summer sports; Canada, the Netherlands and Switzerland for winter sports), i.e. being successful given their expenditure on elite sport, were also the countries that performed best on Pillar 2 -organisation, governance and structure of elite sport- all with scores far above the average.

Another way to look at the results is with a horizontal analysis of Figure 2, considering the performance of the sample nations against all of the Pillars compared with the sample average and the maximum scores on each Pillar, for example, using radar charts. Figure 3 is an example of Japan and Brazil, as two (future) hosts of the summer Olympic Games (at the time of writing this paper). The figure clearly illustrates the contrast of both countries in most Pillars and how elite athletes, coaches and performance directors have evaluated them. The main weakness in Brazil, covering all Pillars, is that there is no clear overall plan, leadership and coordination to

be successful in elite sport in the short term. Without going into depth on the results, this CIs approach enables us to quickly discern the relative strengths and weaknesses of each nation and also gives an indication about what are the most obvious areas for improvement. We refer to De Bosscher et al. (2015) for more details about results of individual nations, pillars and CSFs.

INSERT FIGURE 3 NEAR HERE

What CIs May NOT do, and needs to be informed with further qualitative information

The CI has thus far assisted in gaining insights into those Pillars (and CSFs) that are related to success to a certain extent. It is important to note that the SPLISS study used a mixed methods approach, where the results are built mainly by qualitative analysis. The CIs are only used to assist further qualitative analysis and reporting. Overall, the CI can support policy decisions in terms of identifying key success determinants. The findings in Figure 2, also clearly confirm that there is no generic structural or managerial blueprint for success. High performing countries such as France, Australia and Japan show strengths in different sets of Pillars and, furthermore, each Pillar score is composed of different configurations of CSFs. Smaller or less wealthy countries may find a different set of ingredients that work effectively in their given context. None of these approaches is necessarily right or wrong. Elite sport pathways therefore require a contingency approach, with a model that fits best with the unique situation that a country is placed in. As such, the CIs by themselves do not suffice to inform how each Pillar or CSF is developed in a country. The danger of the CI scores for nations is that they seem to have the characteristic of role models that are easily used as benchmarks for many other countries, thereby ignoring the local contexts. It is unclear under what conditions we can expect best practices to work in other contexts or structures, or indeed which factors or characteristics are reciprocal (Houlihan and Green 2008). Further exploration and deeper qualitative analysis

through interviews, additional document analysis or open ended supplementary questions are needed to understand these contexts more broadly. This is one of the main reasons that the SPLISS project collaborated with local researchers and each CSF is further explained by qualitative data. Still, its attempt to implement a generic model and CSFs in a local context is probably one of the greatest challenges for the SPLISS project. Elite sport is part of an open system, as it is influenced by the social, cultural and economic conditions of the community in which it operates (Chelladurai 2009). This suggests examining policies beyond the formal structures that are in place to develop elite athletes. The CI scores do not provide this information because they are hard to measure and therefore prudence is needed in making simplistic policy conclusions. CIs are, above all, the sum of their parts.

The greatest problem in constructing a CI is the lack of relevant data (Freudenberg 2003). Data or statistics may be unavailable, or may be available but not comparable across countries, or may exist only for a few countries. In SPLISS, this is partly addressed with the surveys. Further descriptive data can help understand the context better, but this is ignored in the CI, which may lead to misleading conclusions when focusing on the scores only. A typical example concerns the focus on factors that are predominantly driven by national governments and national sporting organisations, therefore excluding the more commercial or 'market models' of elite athlete development or regional elite sport development. While this excluded funding can have a significant effect on success, a lack of transnationally comparable data prevents our analysis from including such funding sources. For example, only a few countries have reliable data on what is spent by commercial sponsors or regional departments and local governments on elite sport. In the case of Brazil, for instance, further qualitative analysis through interviews revealed that, in addition to the 147 million euros of nationally coordinated elite sport funding in 2011 (measured in SPLISS) from the federal government and lotteries (through the tax break act, the Piva Act) and the Olympic Committee, an estimated further 75 million euros is provided

by state companies, which are some of the largest sport sponsors in Brazil. While in most countries such entities are commonplace, particularly around infrastructure such as railways and telecommunications, strategic goods and services (mail, weapons), in Brazil, the impact on sport is different because of the considerable amounts of money invested directly into national governing bodies (such as judo, tennis, boxing and volleyball). They have been deliberately excluded from SPLISS because there is no national coordination of state company funding, nor are there clear criteria, which marked out the SPLISS study. As another example, while in absolute terms, Korea stands out as the most substantial investor in elite sport, spending more than 200 million euros per year (at the national level by governments, lotteries and the Olympic Committee), notably 53% of this expenditure is provided for the organisation of international events, such as the Asian Games, the IAAF World Championships and the Universiade, and as such these are related to long-term investments in infrastructure. These cases illustrate that further qualitative analysis is needed to interpret the scores in their context.

As another example, it is difficult to take into account the role of the states and regions in elite sport development, as well as municipalities. In the case of Canada, Australia and France, this is assessed as being under-evaluated in the CSFs, especially in relation to elite sport facilities (Pillar 6). Here, the SPLISS model is limited to how such funding and activities are nationally coordinated, but does not include these measurements as such. Canada probably suffers most from the omission of sub-national data, in particular in relation to Pillar 3, where the states are responsible for grass roots sport development and Sport Canada (the national sport association) for elite sport development; in addition, as sport and physical education in schools is a regional responsibility, data were not available for CSFs on these matters. Despite this, there is only little alignment between the existing regional Canadian Sport Centres and the support services athletes get at different levels.

Finally, a comparison of a small and a large country, treated as equals in a composite indicator, can be deceptive considering their relative size (Freudenberg 2003). Small countries face different problems to large countries. It can be argued that the smaller nations, such as Denmark, the Netherlands and Switzerland, can differentiate themselves from bigger nations in their ability to utilise the potential of their athletes to create elite sport achievements and to coordinate elite sport, with relatively high autonomy given to the sports. It was in this respect remarkable that Denmark, with a population of 5.6 million, could deliver an overall level of performance that is in the same cluster as Brazil, which has a population of around 200 million.

In conclusion, this section illustrated that, despite the theoretical insights that CIs can give to elite sport development, there are a number of drawbacks involved that cannot be ignored. These can in part be addressed by additional qualitative analysis and are therefore a plea to use mixed methods research at different levels of data collection, analysis and interpretation.

Discussion

This paper has attempted to demonstrate the utility of mixed methods research, in international comparative studies on elite sport policies, with a particular focus on the composite indicators. Inspired by economic methods, it showed how a complex and large amount of international data on elite sport policies in 15 nations (over 3000 pages and responses of 3142 elite athletes, 1376 elite coaches and 241 performance directors) have been compared and objectified into easily understood formats. The literature (table 2) has shown that there does not exist a blueprint to develop CIs in one particular way and therefore different conceptualisation methods, scoring methods and standardisation methods should be considered for researchers intending to use CIs. The CI is a helpful tool primarily in identifying possible relationships with success, identifying specific success factors in elite sport policies, facilitating interpretation and comparison, understanding differences and convergences of elite sport

systems, simplifying the visualisation of results, and identifying any specific characteristics in the overall results for the nine Pillars and their relationship with success. These methods can be applied to other sport policy evaluations for example to compare large samples, to extract meaning from qualitative data, and to facilitate pattern recognition (Sandelowski, Voils, and Knafl 2009). But there are also drawbacks, one of the main challenges being to understand elite sport policies as part of a broader social, cultural and political context. Considering that the contribution of CIs and the disadvantages were already discussed in the results section, Table 6 displays a summary of the pros and cons for the evaluation and international comparison of elite sport policies, referring to economic competitiveness studies using CIs (e.g. Freudenberg 2003, Nardo et al. 2008, Berger and Bristow 2009, Maxwell 2010, Sandelowski, Voils, and Knafl 2009) and applied to SPLISS. These elements should be considered when researchers intend to use CIs or mirror to SPLISS to compare policies internationally.

INSERT TABLE 6 NEAR HERE

To overcome these drawbacks, the paper has argued for the need to integrate additional qualitative information in data collection, analysis and reporting to establish general conclusions and address ‘how’ and ‘why’ questions, rather than simply ‘whether’ and ‘to what extent’ (Maxwell 2010). As argued by supporters of mixed methods research, these methods provide strengths that offset the weaknesses of both quantitative and qualitative research (Creswell and Plano Clark 2007, Rudd and Johnson 2010). Research methods in the SPLISS study were mixed at the level of data collection, data analysis and interpretation. SPLISS does not claim to be an all-inclusive and perfectly comparable method. What it strives for, is to use mixed methods to offset weaknesses of both quantitative and qualitative methods, and this way it strives for completeness. The methodology illustrated in this paper, can advance the knowledge

about developing a theory of sports policy factors leading to international sporting success by blending different methods, mainly in two ways.

First, by evaluating the existence of various system components as well as the rating given by athletes, coaches and performance directors, the methods can give a more comprehensive understanding to assessing the "black box" of throughput-elements in elite sport policies that are difficult to evaluate. These data (collected with inventories and surveys) are seen as complementary and help to improve the content and construct validity of the theoretical model (De Bosscher 2016). By including perceptual measures alongside objective ones, scholars may examine the degree to which structures, processes, and outcomes align with the perceptions of those participating in the organization on a day-to-day basis. According to Dellinger and Leech (2007, p716) only in this way can researchers acquire a more comprehensive understanding of any possible organizational dysfunctions that may be undermining the effectiveness of the organization. The other studies in sport management that used CIs (shown in Table 2) did not survey stakeholders, in addition to their qualitative evaluation.

Second, in addition to qualitative methods, measuring policies using quantitative CIs, as an aggregated score of 750 sub-elements, is a useful tool to (a) transcend the descriptive level of comparison and thus to extract meaning from qualitative data and to verify interpretations; (b) facilitate pattern recognition of policies (Sandelowski et al., 2009); and (c) improve criterion validation of the conceptual model by relating inputs and throughputs to outputs. It is however important to note that the scoring system is rather a supportive and tangible way of understanding policies more broadly than an isolated measurement or ranking system by itself. Accordingly, one main criticism given to the use of CIs in competitiveness literature is the absence of a coherent framework capable of providing appropriate guidance on the selection of variables, their relative weights and their inter-relationships (Freudenberg 2003). As a consequence, index crafters in economies rely very heavily on expert judgement and ad hoc

empirical analysis (Berger and Bristow 2009). Policies are often non-quantifiable and hard to measure, which poses a great challenge to ascertaining the validity of performance assessment (Thiel 2002). In order to strive for completeness, the SPLISS model builds on an extensive literature review on elite sport systems and success factors, surveys with elite athletes, interviews with experts and a pilot study in six nations to compose a comprehensive model of nine Pillars and identify appropriate CSFs (De Bosscher et al. 2008, De Bosscher et al. 2006). Still, it is not an all-embracing model that can be applied to any situation, any country or any context. 'Its function is not deterministic: rather it aims to identify pivotal issues and to generate crucial questions in a benchmark study of elite sport systems' (De Bosscher et al. 2006, 209). Therefore, the potential of mixed research methods has not yet been fully realised in this study. Discovering how policy can lead to success is yet another step to take, as there are clearly different paths to success, that are context and sport specific (Truyens et al. 2013, Brouwers, Sotiriadou, and De Bosscher 2014, De Bosscher et al. 2015). The SPLISS model is therefore dynamic, and will continually be adapted over time and to different sport settings, different sport contexts and different situations. As Morgan (2007, p71) suggests, "by moving 'back and forth' between induction and deduction, one can convert observations into theories and then assess those theories through action". Constructing validity "is a continuing process of experimentation and modification leading to the refinement of the instrument that measures the construct" (Morgan 2007, p80)

The results demonstrated in Figures 2 and 3 cannot illuminate the full complexity and richness of sport policy in relation to success, but can provide a deeper understanding of some substantive resources that are mostly similar -key CSFs that are present among high performing countries- and are possibly evidence of convergence. For policy makers and high performance managers, CIs can support decision making and facilitate policy implementation more efficiently and effectively because they can summarise complex, multi-dimensional realities

(Nardo et al. 2008). However, the danger is that, while CIs are easy to interpret and to facilitate communication, they may also send misleading policy messages if poorly constructed or misinterpreted. Misinterpretation is likely if dimensions that are difficult to measure are ignored (Nardo et al. 2008), such as the socio-cultural context of nations, or the specificity of sports. It can become problematic when nations use this information to benchmark themselves and simply transfer best practices, Pillars and critical success to their own context (e.g. Andersen and Ronglan 2012, Böhlke and Robinson 2009, Henry et al. 2005). The discussion of this phenomenon in sport was raised by Houlihan and Green (2008), who borrowed the concepts of 'policy learning', 'lesson drawing' and 'policy transfer' from social policy literature. They state that one of the most significant obstacles in the implementation of policies is the inherent difficulty of translating ideas and strategies from one context into another. The CI measurement may then have intended and unintended effects, and may have a retroactive impact on the policy decision-making process in some nations. The reality in the SPLISS model is that there is no blueprint but a set of broad principles around a common framework that can be adapted to local circumstances in a culturally appropriate manner. Accordingly, the key challenge for nations is still to find the right blend of system ingredients and processes that work best in their own context and culture, encouraging them to "benchlearn" from rivals rather than merely benchmarking against them (De Bosscher et al. 2015). The key point of note is that the radar graphs and traffic lights are not a stand-alone evaluation of elite sport policies; they cannot be isolated from the general descriptive information on elite sport policies. Essentially, qualitative and quantitative data remain complementary.

Finally, the main drawback using mixed methods research is probably that the methods as they are used in SPLISS are very time consuming and this, in combination with the fact that the use of surveys in an international context is expensive, makes comparative mixed research methods studies very labour intensive. It is challenging to manage data from over 3000 pages of

(mainly qualitative) inventory information from 15 nations, combined with large sample survey data, and to compose CSF scores that are further aggregated into CIs. Applying these methods to a larger sample of nations (e.g. like in competitiveness studies) would require data reduction and selection, with the danger of losing the holistic approach to elite sport policies. The fewer the number of performance indicators there are, the more difficult it becomes to obtain an accurate report of the performance (Thiel 2002). Furthermore, such a process, from the country selection to the final CIs and mixed methods report, easily takes three years. As a consequence, practice sometimes develops faster than theory.

References

Andersen, Svein, Barrie Houlihan, and Lars Tore Ronglan. 2015. *Managing elite sport systems. Research and practice*. New York: Routledge.

Andersen, Svein, and Lars Tore Ronglan. 2012. *Nordic Elite Sport: Same Ambitions, Different Tracks*: Copenhagen Business School Press.

Berger, Thomas, and Gillian Bristow. 2009. "Competitiveness and the Benchmarking of Nations—A Critical Reflection." *International Advances in Economic Research* 15 (4):378-392. doi: 10.1007/s11294-009-9231-x.

Böhlke, Nikolai, and Leigh Robinson. 2009. "Benchmarking of élite sport systems." *Management Decision* 47 (1):67-84.

Brouwers, Jessie, Popi Sotiriadou, and Veerle De Bosscher. 2014. "Sport-specific policies and factors that influence international success: The case of tennis." *Sport Management Review* (0). doi: 10.1016/j.smr.2014.10.003.

Chelladurai, P. 2001. *Managing organisations for sport and physical activity: A systems perspective*. Scottsdale, AZ: Holcomb Hathaway Publishers.

Chelladurai, P. 2009. *Managing organisations for sport and physical activity: A systems perspective*. Arizona: Holcomb Hathaway Publishers.

Creswell, J.W. 2009. *Research design: Qualitative, quantitative, and mixed research methods approaches*. Los Angeles: Sage Publications.

Creswell, J.W, and V.L Plano Clark. 2007. *Designing and conducting mixed methods research*: SAGE Publications.

De Bosscher, V. 2007. *Sports policy factors leading to international sporting success*. Edited by Published doctoral thesis. Published doctoral thesis. Brussels: VUBPRESS.

De Bosscher, V. 2016. "Theory of Sports Policy factors Leading to International Sporting Success (SPLISS)." In *Handbook of Theory in Sport Management*, edited by B.C. Cunningham, J.S. Fink and A. Doherty, 93-109. London: Routledge.

De Bosscher, V., J. Bingham, S. Shibli, M. Van Bottenburg, and P. De Knop. 2008. *The global Sporting Arms Race. An international comparative study on sports policy factors leading to international sporting success.* Aachen: Meyer & Meyer.

De Bosscher, V., P. De Knop, and M. van Bottenburg. 2009. "An analysis of homogeneity and heterogeneity of elite sports systems in six nations." *International journal of sports marketing & sponsorship* 10 (2):111-131.

De Bosscher, V., P. De Knop, M. Van Bottenburg, and S. Shibli. 2006. "A Conceptual Framework for Analysing Sports Policy Factors Leading to International Sporting Success." *European sport management quarterly* 6 (2):185-215.

De Bosscher, V., S. Shibli, M. van Bottenburg, P. De Knop, and J. Truyens. 2010. "Developing a methodology for comparing the elite sport systems and policies of nations: A mixed research methods approach." *Journal of Sport Management* 24:467-600.

De Bosscher, V., S. Shibli, H. Westerbeek, and M. van Bottenburg. 2015. *Successful elite sport policies. An international comparison of the Sports Policy factors Leading to International Sporting Success (SPLISS 2.0) in 15 nations.* Aachen: Meyer & Meyer.

De Pelsmacker, P, and P Van Kenhove. 1999. *Marktonderzoek: methoden en toepassingen [Market research: methods and applications]* 3rd ed. Leuven Apeldoorn, BE: Garant.

Dellinger, Amy B., and Nancy L. Leech. 2007. "Toward a Unified Validation Framework in Mixed Methods Research." *Journal of Mixed Methods Research* 1 (4):309-332. doi: 10.1177/1558689807306147.

Fraser Institute. 2005. Economic Freedom of the World. Annual Report. Vancouver: Fraser Institute.

Freudenberg, Michael. 2003. "Composite Indicators of Country Performance: A Critical Assessment." *OECD Science, Technology and Industry Working Papers* 16. doi: 10.1787/405566708255.

Henry, Ian, Mahfoud Amara, Mansour Al-Tauqi, and Ping Chao Lee. 2005. "A typology of approaches to comparative analysis of sports policy." *Journal of Sport Management* 19 (4):480-496.

Houlihan, B. 2013. "Commercial, political, social and cultural factors impacting on the management of high performance sport." In *Managing high Performance Sport*, edited by Popi Sotiriadou and Veerle De Bosscher, 17-27. London and New York, UK: Routledge.

Houlihan, B., and M. Green. 2008. *Comparative Elite Sport Development. Systems, structures and public policy*. London, UK: Elsevier.

Institute for Management Development (IMD). 2012. *The World Competitiveness Yearbook 2012*. Lausanne: : Thomson Learning.

Johnson, R. Burke, and Anthony J. Onwuegbuzie. 2004. "Mixed methods research: A research paradigm whose time has come." *Educational Researcher* 33 (7):14-26. doi: 10.3102/0013189x033007014.

Linssen, G. W. J. M. 1998. "Benchmarking. De concurrentietoets 1997: Een voorbeeld van benchmarken [An example of benchmarking]." *Beleidsanalyse* 1:14-22.

Maxwell, Joseph A. 2010. "Using numbers in qualitative research." *Qualitative Inquiry* 16 (6):475-482. doi: 10.1177/1077800410364740.

Morgan, David L. 2007. "Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods." *Journal of Mixed Methods Research* 1 (1):48-76. doi: 10.1177/2345678906292462.

Nardo, M., M. Saisana, A. Saltelli, S. Tarantola, A. Hoffman, E. Giovannini, and OECD. 2008. Handbook on Constructing Composite Indicators: Methodology and user guide. OECD report.: OECD.

Ochel, W, and Oliver Röhn. 2006. "Ranking of Countries - the WEF, IMD, Fraser and Heritage Indices." CESifo DICE Report 4 (2):48.

Onsel, Ulengin, Ulusoy, Aktas, Kabak, and Topcu. 2008. "A new perspective on the competitiveness of nations." *Socio-Economic Planning Science* 42:221-246.

Rudd, Andy, and R. Burke Johnson. 2010. "A call for more mixed methods in sport management research." *Sport Management Review* 13 (1):14-24. doi: 10.1016/j.smr.2009.06.004.

Sandelowski, Margarete, Corrine I. Voils, and George Knafl. 2009. "On quantitizing." *Journal of Mixed Methods Research* 3 (3):208-222. doi: 10.1177/1558689809334210.

Shibli, S. 2003. "Analysing performance at the Olympic Games: Beyond the final medal table." 11th Congress of the European Association for Sport Management, Stockholm, Sweden.

Shilbury, D., and Kathleen A. Moore. 2006. "A Study of Organizational Effectiveness for National Olympic Sporting Organizations." *Nonprofit and Voluntary Sector Quarterly* 35 (1):5-38. doi: 10.1177/0899764005279512.

Tashakkori, A., and C. Teddlie. 2003. *Handbook of Mixed Methods in Social & Behavioral Research*. Thousand Oakes, CA: SAGE Publications.

Thiel, S.V. and Leeuw, F.L. 2002. "The performance paradox in the public sector. ." *Public performance & management review* 25 (3): 267-281.

Truyens, Jasper, Veerle De Bosscher, Bruno Heyndels, and Hans Westerbeek. 2013. "A resource-based perspective on countries' competitive advantage in elite athletics." *International Journal of Sport Policy and Politics* 6 (3):459-489. doi: 10.1080/19406940.2013.839954.

Truyens, Jasper, Veerle De Bosscher, Popi Sotiriadou, Bruno Heyndels, and Hans Westerbeek. 2015. "A Method to Evaluate Countries' Resource-Based Competitive Advantage. A Case Study in Athletics." *Sport Management Review*. doi: http://dx.doi.org/10.1016/j.smr.2015.05.002.

van der Roest, Jan-Willem, Ramón Spaaij, and Maarten van Bottenburg. 2015. "Mixed Methods in Emerging Academic Subdisciplines: The Case of Sport Management." *Journal of Mixed Methods Research* 9 (1):70-90. doi: 10.1177/1558689813508225.

Vrkljan, Brenda H. 2009. "Constructing a Mixed Methods Design to Explore the Older Driver—Copilot Relationship." *Journal of Mixed Methods Research* 3 (4):371-385. doi: 10.1177/1558689809336843.

World Economic Forum (WEF). 2007. The global competitiveness report 2007. New York: Palgrave Macmillan.

World Factbook. 2012. "World Factbook." [Internet source] Accessed March, 3rd. https://www.cia.gov/library/publications/the-world-factbook/.

Acknowledgement

The coordination of the SPLISS research was funded by the Vrije Universiteit Brussel

SPLISS has largely depended on institutional investments to ensure the national coordination of this large-scale project. The SPLISS consortium partners did not seek any financial contribution from the participating nations as a return for the massive coordinating work. All such costs were met by the Vrije Universiteit Brussels and supported by the consortium members and engaged researchers from the partner countries. The author would like to thank the policy institutions that funded the national data collection (see De Bosscher et al. 2015, p24-29 for an overview)

The authors would like to thank the dedicated researchers partners involved in the SPLISS-2.0 study, among others (sorted by country): Hans Westerbeek and Camilla Brockett (Australia, Victoria University) ● Stephanie De Croock and Jasper Truyens (Belgium (Fla), Vrije Universiteit Brussel) Mathieu Winand (Stirling University) and Thierry Zintz (Belgium (Wal), Université catholique de Louvain) ● Maria Tereza Silveira Bohme & team (Brazil, University of São Paulo) ● David Legg & team (Canada, Mount Royal University) ● Henrik Brandt, Rasmus K. Storm, Lau Tofft and Nynne Mortensen (Denmark, Danish Institute for Sports Studies & University of Southern Denmark) ● Eerik Hanni (Estonia, National Audit Office of Estonia) ● Patrick Mignon and Emanuel Lelore (France, Institut National du Sport et de l'Éducation Physique-INSEP) ● Jari Lämsä, Jarmo Mäkinen and Mikko Kärmeniemi (Finland, KIHU - Research institute for Olympic Sports) ● Yoshiyuki Mano, Hiroaki Funahashi and team (Japan, Waseda University) ● Maarten van Bottenburg and Bake Dijk (the Netherlands, Utrecht University) ● Simon Shibli (Sheffield Hallam University) and Paul Donnelly (Sport Northern Ireland, UK) ● Pedro Guedes De Carvalho and Rui Canelas (Portugal, Beira Interior University) ● Anna Vilanova, Eduard Inglés and team (Spain, National Institute of Physical Education of Catalonia, INEFC) ● Eunha Koh (South-Korea, Korea Institute of Sport Science)● Hippolyt Kempf, Marco Stopper & Andreas, Christophe Weber (Switzerland, Swiss Federal Institute of Sport Magglingen SFISM).

Footnotes

[1] Flanders is the northern, Dutch speaking part of Belgium (6.3 million inhabitants), Wallonia the southern, French and German speaking part (4.0 million inhabitants). In Belgium the Flemish community (Flanders) and the French/German speaking community (Wallonia) have separate sport policies at each level, from local to national (including three separate ministers of sport). Apart from the Olympic Committee (BOIC), whose main task is to select athletes for the Olympic Games, there is no national (federal) policy or structure for sport, nor are there expenditures on sport at federal level. Therefore Flanders and Wallonia are seen in this research as if it is two distinct nations. It was an established fact that policy analysis for Belgium as a nation could not be determined by summing both regions.

For Northern Ireland, UK Sport is the coordinating authority for elite sport, where DCAL (government department for culture, media and sport) in Northern Ireland sets the policy direction and Sport NI puts this into practice. Some sports are supported at UK-level, others are supported at the home nation level of Northern Ireland.

[2] An Elite athlete was defined as (1) an (able bodied) athlete who, whether as an individual, or as part of a team, is ranked in the world top 16 for his or her discipline, or in the top 12 of any equivalent Continental ranking system." OR (2) "An athlete who receives direct or indirect funding and/or other services via a support programme funded and/or organised on a national (or regional) basis for the purpose of achieving success at least one of the following levels: the Olympic Games; the senior World Championships; and the senior Continental Championships in his or her sport (European, Asian, Pan American etc).

[3] Market share is a standardised measure of total achievement in an event whereby total medals won are converted into 'points' (gold=3, silver=2, bronze=1) and the points won by a given nation is subsequently expressed as a percentage of the total points awarded (Shibli 2003)

TABLES

Table 1: Overview of international comparative elite sport policy books (ranked chronologically)

Authors	**Methods & study characteristics**	**FOCUS: Countries & sports**
Green & Houlihan, 2005	QUALITATIVE - descriptive and comparative; qualitative data collection: interviews & document analysis - advocacy coalition framework (ACF) to analyze policy changes - aims to identify the degree of similarity in elite sport development models	3 countries: Australia, UK and Canada overall + 3 sports: swimming, athletics, sailing/ yachting.
Digel et al. (2006)	QUALITATIVE - descriptive and comparative; - qualitative data collection: interviews & document analysis; questionnaires (3 sport associations, Ministries responsible for HP sport, National Olympic committees, country specific organisations)	8 countries: Australia, China, France, Germany, Italy, Russia, UK, USA Overall + 3 sports: Athletics, Swimming, Volleyball
Bergsgard et al. (2007)	QUALITATIVE - textbook: contributions from researchers of each country: author(s) from each country describe elite sport policies, generally based on literature & document analysis - contextual & historical - descriptive;	4 countries: Germany, England, Canada, Norway Overall (national) level
Houlihan & Green (2008)	QUALITATIVE - textbook: contributions from researchers of each country: author(s) from each country describe elite sport policies, generally based on literature & document analysis - contextual & historical; builds on Green and Houlihan's (2005) ACF - descriptive;	9 countries: China Japan, Singapore, Germany, France, Poland, Norway, New Zealand, USA Overall (national) sports level
De Bosscher et al. (2008)	SPLISS 1.0 MIXED METHODS: sequential qualitative phase 1 (model); concurrent qualitative + quantitative phase 2 (international comparison) - meso-level factors only: SPLISS framework - objective inventory + stakeholder involvement (1090 athletes, 273 coaches, 71 performance directors) - scoring system & traffic lights: to move beyond the descriptive level of comparison	6 countries: Belgium (Flanders + Wallonia), Canada, Italy, The Netherlands, Norway, UK Overall (national) sports level

Andersen & Tore Ronglan (2012)	QUALITATIVE - textbook: contributions from researchers of each country: - contextual & historical; aims to look beyond the designated sport organizations & to capture the processes behind national elite sports systems; aims to identify how national elite sport systems came about, how they relate to success in individual sports, how they differ in terms of centralization, responsibilities and roles, and how this influences the capacity for successful elite sport development	4 Nordic countries (with similar systems) + success stories of sports: Norway (women's handball), Finland (men's ice hockey), Denmark (track cycling), Sweden (tennis & golf);
De Bosscher et al. (2015)	SPLISS 2.0 – follow up building on the methods of SPLISS 1.0 MIXED METHODS: concurrent qualitative + quantitative (international comparison) - meso-level factors only: SPLISS framework - objective inventory + stakeholder involvement (3142 athletes, 1376 coaches, 243 performance directors) - composite indicators	15 nations: 10 European, 2 American, 2 Asian and Australia.

Table 2: Comparison of composite indicators used in international comparative studies in the field of economy and sport management/policy (ordered chronologically)

	ECONOMIC COMPOSIT INDICATORS (CI)				SPORT MANAGEMENT/POLICY COMPOSIT INDICATORS (CI)				
Measurement	**WEF, 2012 GCI - Global Competitiveness Report**	**IMD, 2015 World competitiveness yearbook**	**EFW - Fraser Institute, 2005**	**Heritage Foundation Index of Economic Freedom**	**Madella, 2002**	**SPLISS (De Bosscher 2007; 2010)**	**RAT Robinson & Minikin, 2012**	**ORFOC Truyens et al., 2015**	**SPLISS 2.0 De Bosscher et al. (2015)**
Objective	Tries to measure national competitiveness; *the ability of countries to attain sustained economic growth.*	Ranking the ability of nations to create and maintain an environment that sustains the competitiveness of enterprises and promotes economic growth	Measures the degree of economic freedom present in five major areas	Systematic, empirical measurement of economic freedom	Measure the performance of four national swimming Federations	Gain insights in the key factors in elite sport policies and their relationship with international sporting success (i.e. competitiveness of elite sport at the meso-level);	Evaluate the competitive advantage of NFs within 3 pacific countries	Method to evaluate the organizational resources and capabilities in athletics	Better understand which (and how) sport policies influence international sporting success; obtain a better insight into the effectiveness and efficiency of elite sport policies of nations at an overall sports level; follow up to SPLISS 1.0
Theory/Concept	3 component indices, calculated on the basis of 35 sub-indices and 131 sub indicators (=unweighted average of data if weights are not given)	4 main competitiveness factors (Economic performance, government efficiency, business efficiency, infrastructure) that are broken down into 5 subfactors. 333 competitiveness Criteria	21 components in five major areas are incorporated into the index; made up of several sub-components	50 independent variables divided into 10 broad factors of economic freedom	Seven organizational dimensions with 32 quantitative criteria	SPLISS framework: 9 Pillars at the input-throughput-output levels; operationalised into 103 Critical Success factors (CSF) and 41 that were not included in the study	RAT model: competitive scores based on eight Pillars of performance	ORFOC-framework: 98 organisational resources and first-order capabilities in athletics; reflecting the nine Pillars of the SPLISS model	SPLISS framework: 9 Pillars at the input-throughput-output levels; operationalised into 96 Critical Success factors and 750 sub-factors
Nations	131	61 nations (9 regions)	-	155	4	6	3	4	15
Data collection	QUANTITATIVE 14 hard data and 21 survey data (e.g. in 2004 2,100 executives from 30 OECD member countries)	QUANTITATIVE 128 hard data and 113 survey data (e.g. in 2005, 4000 executives from 60 economies); 2/3 statistical data, 1/3 survey data	QUANTITATIVE 19 hard data and 19 survey data supplied by WEF and IMD surveys	QUANTITATIVE	QUANTITATIVE Hard quantitative data on 32 organisational indicators	MIXED METHODS Inventory with qualitative and quantitative data: 54 hard data combined with elite sport climate survey: 18 hard survey data and	QUANTITATIVE A self-assessment tool (RAT) to measure the development level of organisational resources of NFs	MIXED METHODS Resource inventory to collect quantitative and qualitative information delivered by local researchers	MIXED METHODS Inventory with qualitative and quantitative data (212 questions), collected from policy documents and interviews; elite

						31 perceived survey data, from 1090 athletes, 273 coaches and 71 performance directors; Outputs: top 3 OG, WC, 1 year			sport climate survey objective and perceived data with 3142 elite athletes, 1376 coaches & 243 performance directors Outputs: top 3-8 OG, WC, 4 years (Infostrada database)
Stakeholders involvement/ surveys	YES	YES	YES	YES	NO	YES	NO	NO	YES
Data analysis	*QUANTITATIVE*	*QUANTITATIVE*	*QUANTITATIVE*	*QUANTITATIVE*	*QUANTITATIVE*	MIXED METHODS	MIXED METHODS	MIXED METHODS	MIXED METHODS
Standardization & Scoring	*(1)Score classes method:* scale with scores ranging from 1- 7 (based on linear interpolations to normalize the indicators within a scale). (2*) continuous scaling methods* transform the underlying indicator values into a continuous, uniform scale that retains the relative distances between the original values. The distance from the best and worst performer to transform the original indicators into a range between 0 and 1. In a second step, GCI linearly transform these values to lie within a range of 1 and 7 (Min-Max method)	*Continuous scaling method:* - transforms all original indicators into a common scale. Converts data from a 1-6 scale to a 0-10 scale and then calculates standard deviation values to determine rankings (Rosselet, 2008). - slightly different linear interpolation. All original indicator values are transformed into a standardized distribution with mean 0 and standard deviation of 1 (Ochel, Röhen, 2006)	*(1) Score classes method:* scale with scores ranging from 1- 10 (based on linear interpolations to normalize the indicators within a scale). (2) *continuous scaling methods* transform the underlying indicator values into a continuous, uniform scale that retains the relative distances between the original values. The distance from the best and worst performer to transform the original indicators into a range between 0 and 1. In a second step, EFW linearly transform these values to lie within a range of 0 and 10. (Min-Max)	*Score classes method:* The 50 independent variables are analyzed to determine for each of the 10 factors a score on a scale running from 1 to 5. *Score classes method* (or categorical scaling method), using expert assessments to determine final score	*Scores* for indicators on seven dimensions are standardised based on total population, GDP, average sport participation, number of swimming pools, and relative positioning of swimming compared to other sports	CI used only to inform qualitative interpretation *Score classes method 1-5:* - Qualitative data transformed in scores - scale with scores ranging from 1-5 for each CSF are aggregated into one final percentage score for each Pillar. - These total scores are categorized into five scales (traffic lights) - There are no overall rankings (or indices) made and Pillars are not aggregated into one final score - The allocation of scores is for qualitative data (overall sports policy inventory) uses expert assessments to determine final score - objective and	*Score classes method:* Pillar scores for specific NFs were averaged for an NOC on a 0–4 scale (no development – professionalised and specialised level of development)	*Score classes method:* scale between 0 - 1: - qualitative indicators as dummy variables - nominal scales (0-1) of different sub factors - quantitative data: distance from the group leader Total CI score is the weighted average of it's indicators.	CI used only to inform qualitative interpretation *Score classes method:* - Qualitative data transformed in scores - 750 sub-factors are scored on a scale between 0 – 1, depending on the kind of question: (a) inventory quantitative: standardized z-scores, distance from mean, cumulative probability score (b) inventory qualitative: aggregated sum of dichotomous qualitative variables; uses expert assessments to determine final score (c) survey objective: absolute standards (d) survey perceived (likert): weighted ratings

						subjective information is deliberately kept separate			- subfactors are aggregated into 96 CSFs and subsequently aggregated into one final percentage score - objective and subjective information is merged
Weighting	YES Statistical techniques to determine weights: regression analysis with the average growth rate as the dependent variable to establish the weights of its three subcomponents as well as the weights within these subcomponents.	YES In the first step, the universe of basic indicators are grouped into 20 sub-indices . In the second step all sub-indicators are assigned equal weights in the composite index. *Sub-factors do not necessarily include the same number of criteria*	EQUAL WEIGHTS In the first step, the universe of basic indicators are grouped into 5 sub-indices. In the second step all sub-indicators are assigned equal weights in the composite index	EQUAL WEIGHTS In the first step, the universe of basic indicators are grouped into 10 sub-indices. The 10 factors are weighted equally. In the second step all sub-indicators are assigned equal weights in the composite index	NO	YES Weightings are allocated to aggregate CSFs into one final percentage score according to their relative importance (expert opinion) and to lock in the impact of each CSF on the overall scores. Pillar scores are not merged into one overall score	NO	YES Weighted Index Scores (WISs): determined by an expert group rating the variables from 1 (basic value) to 3 (high value)	YES Weightings are allocated to aggregate CSFs into one final percentage score according to their relative importance (expert opinion) and to lock in the impact of each CSF on the overall scores. Pillar scores are not merged into one overall score

Note: some economic indices are produced yearly and the number of countries can differ by year.

WEF: World Economic Forum – GCI: Global Competitiveness Index from the Global Competitiveness Report; IMD: Institute for Management Development; EFW: Economic Freedom of the World index (Fraser Institute); SPLISS: Sports Policy factors Leading to International Sporting Success; NF: national federations; NOCs: National Olympic Committee; RAT: Readiness Assessment Tool; OG: Olympic Games; WC: World Championships.

Table 3: Overview of the number of critical success factors (CSFs) in the nine Pillars measured in SPLISS 2.0 (De Bosscher et al. 2015)

	CSF	Sub-factors
Pillar 1: Financial support	8	9
Pillar 2: Governance, organisation & structure	18	119
Pillar 3: Sports participation	10	31
Pillar 4: Talent identification & development	12	169
Pillar 5: Athletic and post athletic career support	7	122
Pillar 6: Training facilities	9	84
Pillar 7: Coach provision & development	16	100
Pillar 8 (Inter)national competition	7	51
Pillar 9: Scientific research & innovation	9	65
TOTAL	**96**	**750**

Table 4: Overview of SPLISS 2.0 nations clustered according to population and GDP per capita

Nations with a population	**Population**	**GDP/cap**
< 15 million		**(PPP)**
Portugal	10,813,834	$ 22,900
Belgium	10,449,361	$ 37,800
Flanders	6,367,963	
Wallonia (incl. Brussels)	4,081,398	
Switzerland	8,061,516	$ 54,800
Denmark	5,569,077	$ 37,800
Finland	5,268,799	$ 35,900
Northern Ireland (UK)	1,810,863	$ 36.700
Estonia	1,257,921	$22,400
15-40 million		
Canada	34,834,841	$ 43,100
Australia	22,507,617	$ 43,000
Netherlands	16,877,351	$ 43,300
> 40 million		
Brazil	202,656,788	$ 12,100
Japan	127,103,388	$ 37,100
France	66,259,012	$ 35,700
South-Korea	49,039,986	$ 33,200
Spain	47,737,941	$ 30,100
Total SPLISS sample	**620,697,656**	

Source: World Factbook (2012)

Table 5: Spearman's Rank correlations (r^2) for nine Pillars with success (market share summer and winter sports, 2009–2012)(De Bosscher et al. 2015)

	r_s summer	Sig	r_s winter	Sig	N
Pillar 1 (financial support)	**0.909****	0.000	**0.588***	0.039	16
Pillar 2 (organisation & structure)	**0.720****	0.004	**0.685****	0.007	14
Pillar 3 (sport participation)	0.049	0.873	0.267	0.377	13
Pillar 4 (talent ID/TD)	-0.148	0.707	0.237	0.435	13
Pillar 5 (post) athletic career support	0.483	0.080	0.322	0.261	14
Pillar 6 (facilities)	**0.704****	0.005	0.354	0.214	14
Pillar 7 (coaches)	**0.606***	0.028	**0.779****	0.002	13
Pillar 8 (inter)national competition	**0.577***	0.039	0.271	0.370	13
Pillar 9 (scientific research)	**0.71****	0.004	**0.784****	0.001	14

*Note: ** P<0.01; * p<0.05; the correlations are taken only for the countries where data are complete, which explains the different N-values*

Table 6: Pros and cons of using CIs to compare elite sport policies in SPLISS

Pros of using CIs to compare elite sport policies	**Dangers of using CIs to compare elite sport policies**
- Can summarise complex, multi-dimensional realities on elite sport policies into easily understood formats; can enable users to compare complex dimensions effectively. - Useful to facilitate pattern recognition or otherwise to extract meaning from qualitative data, account for all data, document analytic moves, and verify interpretations. - Are easier to interpret than a battery of many separate indicators or descriptive texts; can facilitate communication with general public and promote accountability. - Increase insights into the relationship between elite sport policy indicators and success; between different Pillars and CSFs; between policies and macro determinants (i.e. evaluation of criterion validity). - Comprehensive theoretical construct and CSFs that are transparent. - Larger samples can be compared,	- May send misleading policy messages or be misinterpreted if only looking at total Pillar scores and not how they were constructed. May lead to greater generality for the conclusions than is justified, by neglecting the context. - Overlook how elite sport policies interact with the broader social, cultural and political context. - May invite simplistic policy conclusions. - May be misused, e.g. for benchmarking, with poor policy implementation or to support a desired policy. - May look at the SPLISS model without seeing the social, cultural and economic conditions of the community in which it operates. Danger of making claims about causality that are context specific. - Information is lost in the aggregation process or not measured due to a lack of available and comparable data.

contributes to internal generalizability; enable to characterize the diversity of the group studied - Are reproducible by others. - Can be used to assess progress/change of elite sport policy indicators of countries over time; give an overview of clear action points. - Can help to adequately present evidence for your interpretations and to counter claims - Reduce the visible size of a set of indicators without dropping the underlying information base.	- Compromises made in data collection to avoid comparing apples with oranges may lead to inappropriate policies. - Small and large countries treated as equals can be misleading considering their relative size. - The selection of indicators and weights could be the subject of political dispute. - Time consuming and difficulty to manage all data; time-lag in different countries when reporting data.

FIGURE CAPTIONS

Figure 1: Visual diagram of the (simplified) multilevel SPLISS mixed methods research

design (based on Creswell and Plano Clark 2007)

Figure 2: Success of countries and their scores on the nine elite sport policy Pillars (De Bosscher et al. 2015)

Figure 3: Example of a radar graph of Brazil compared to Japan, compared to the average and maximum scores of 15 nations (De Bosscher et al. 2015)

Figure 1

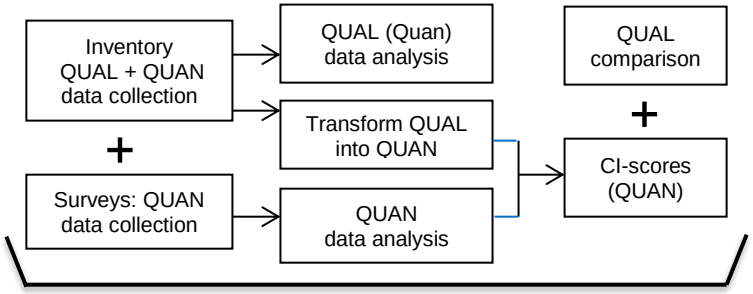
QUALITATIVE
A conceptual model of the Sports Policy factors Leading to International Sporting Success and CSFs

QUALITATIVE + QUANTITATIVE
Test the model in an empirical environment of six nations.

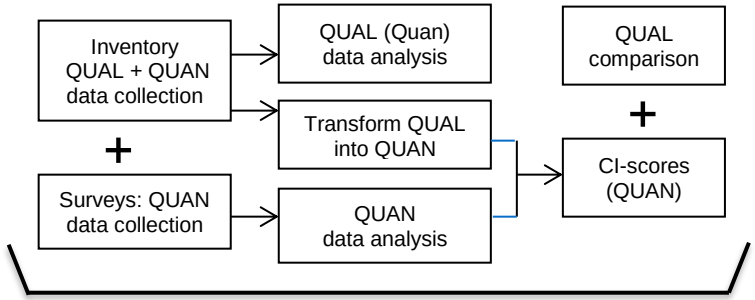
- Improvement in content/construct/ criterion validity of the model
- Measurement of 103 CSFs (41 not measured)
- Insights into the relationship between policies and success, and into elite sport systems of nations in nine Pillars

QUALITATIVE + QUANTITATIVE
Further validation of the nine Pillars and CSFs, comparing 15 nations.

- Reduction to 96 CSFs
- More detailed data collection
- Progressed method to compose CIs in order to compare nations and evaluate effectiveness
- Deeper insights into elite sport systems and the relationship with success between Pillars and CSFs



Embedded concurrent triangulation design (phase 2) (with data transformation)



Embedded concurrent triangulation design (phase 3) (with data transformation)

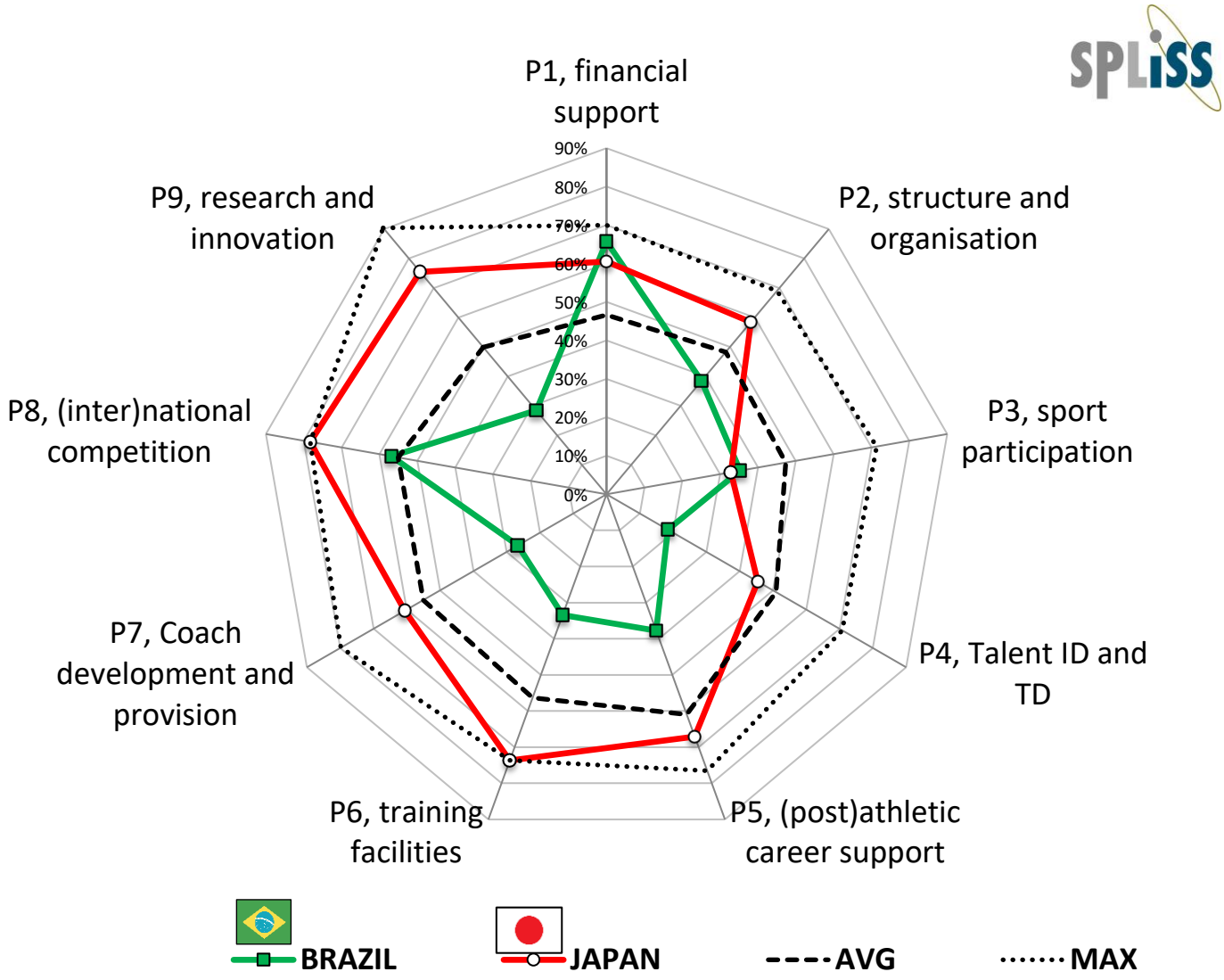
Sequential exploratory design (phases 1–2)

Sequential exploratory design (phases 1–2-3)

Figure 2

Summer	Winter		P1 - Financial support	P2 - Governance, structure & organisation	P3 - Sports participation	P4 - TalentID & developmen	P5 - Athletic career & post career	P6 - Training facilities	P7 - Coach education & provision	P8 - (inter) national competition	P9 - Scientific research & innovation	Average
4.30%	4.38%	**FRA***	69%	*(37%)*	52%	*(45%)*	*(66%)*	*(72%)*	*(80%)*	*(55%)*	*(60%)*	*(60%)*
4.10%	1.21%	**AUS**	60%	64%	54%	49%	76%	66%	69%	48%	90%	64%
3.90%	1.96%	**JAP**	61%	58%	33%	45%	67%	74%	61%	78%	75%	61%
2.40%	6.59%	**KOR***	70%	47%	38%	*(54%)*	54%	55%	*(60%)*	*(57%)*	59%	*(55%)*
1.80%	4.83%	**NED**	45%	69%	62%	68%	77%	65%	62%	54%	53%	62%
1.70%	0.00%	**ESP**	56%	50%	33%	55%	76%	74%	56%	67%	37%	56%
1.50%	12.27%	**CAN**	55%	58%	43%	23%	65%	63%	73%	62%	68%	57%
1.40%	0.00%	**BRA**	66%	38%	35%	18%	38%	33%	27%	57%	28%	38%
0.70%	0.09%	**DEN**	28%	53%	71%	61%	63%	49%	48%	63%	47%	54%
0.60%	3.22%	**SUI**	45%	58%	62%	70%	58%	61%	68%	44%	49%	57%
0.30%	2.52%	**FIN**	36%	47%	50%	49%	70%	43%	56%	65%	53%	52%
0.25%	0.00%	**N-IRL**	30%	42%	42%	41%	63%	60%	53%	40%	31%	45%
0.20%	0.10%	**EST***	26%	*(34%)*	NA	*(64%)*	*(34%)*	*(56%)*	*(34%)*	*(48%)*	*(38%)*	*(42%)*
0.20%	0.19%	**FLA**	41%	47%	48%	71%	66%	47%	52%	45%	52%	52%
0.15%	0.00%	**POR**	25%	34%	41%	44%	49%	48%	52%	52%	35%	42%
0.20%	0.00%	**WAL**	33%	36%	46%	59%	54%	37%	38%	44%	23%	41%
		Average	47%	48%	47%	51%	61%	56%	56%	55%	50%	

Figure 3


P1, financial support
P2, structure and organisation
P3, sport participation
P4, Talent ID and TD
P5, (post)athletic career support
P6, training facilities
P7, Coach development and provision
P8, (inter)national competition
P9, research and innovation
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
BRAZIL
JAPAN
AVG
MAX