# CLUSTERING OF NATIONAL ELITE SPORT POLICY SYSTEMS AND ASSOCIATION WITH INTERNATIONAL SPORTING PERFORMANCE

Anne Renaud[1]& Veerle De Bosscher[2] & Hippolyt Kempf[3]

[1] *Sport Economics Division, Swiss Federal Institute of Sport Magglingen, Magglingen, Switzerland, anne.renaud@baspo.admin.ch*
[2] *Sport and Society Research group, Department Movement and Sport Sciences, Vrije Universiteit Brussel, Brussels, Belgium, veerle.de.bosscher@vub.be*
[3] *Sport Economics Division, Swiss Federal Institute of Sport Magglingen, Magglingen, Switzerland, hippolyt.kempf@baspo.admin.ch*

**Résumé.** Les gouvernements nationaux cherchent des moyens efficaces pour améliorer leurs résultats lors des manifestations sportives internationales telles que les Jeux olympiques. Le projet SPLISS 2.0 (Sport Policy factors leading to International Sporting Success) a observé une grande diversité entre les systèmes de politique du sport d'élite de seize nations. La présente étude a repris les données du projet SPLISS 2.0 afin d'évaluer l'existence de groupes de systèmes de politique du sport similaires. Les performances sportives ainsi que la taille et la richesse des nations ont ensuite été comparées entre les groupes. Une analyse en grappes (*cluster analysis,* hiérarchique et *k-means*) et des tests de *Kruskal-Wallis* et de *Wilcoxon* ont été appliqués pour répondre aux questions d'étude. Quatre groupes de systèmes de politique du sport ont été identifiés (Leading, Challenging, Emerging et Specific). La taille de la population, le PIB par habitant et la performance sportive dans les sports d'été étaient significativement différents entre les groupes; pas les performances pour les sports d'hiver.

**Mots-clés.** Performance sportive, systèmes de politique du sport d'élite, analyse en grappes

**Abstract.** National governments search for efficient ways to improve their results at international sporting events such as the Olympic Games. The SPLISS 2.0 project (Sport Policy factors leading to International Sporting Success) observed a great diversity between the elite sport policy systems of sixteen nations. The current study built on data from the SPLISS 2.0 project to assess the existence of groups of similar sport policy systems. Sports performances as well as the size and wealth of nations were then compared between the groups. Cluster analysis (hierarchical and *k-means*) and *Kruskal-Wallis* and *Wilcoxon* tests were applied to answer the research questions. Four groups of sport policy systems were identified (Leading, Challenging, Emerging and Specific). Population size, GDP per capita and sport performance in summer sports were significantly different between clusters; performance in winter sports wasn't.

**Keywords.** Sporting performance, elite sport policies systems, cluster analysis

## 1 Introduction

The quest for international sporting success such as medals in the Olympic Games has intensified over the last decades. Several factors such as the socio-economic characteristics of nations, national sport policies or the close environment of athletes have been identified as determinants of nations' sport performance (e.g. Bernard and Busse, 2004; Reiche, 2016; Renaud *et al.*, 2018). However, little research has been done on the interplay between sport policy systems, socio-economic characteristics and sporting performance.

Nine elite sport policy areas (*pillars*) that can be developed by national governments were defined in the SPLISS project (Sport Policy factors leading to International Sporting Success). The

SPLISS 2.0 data collection with sixteen nations revealed a high diversity between the elite sport policy systems and indicated that successful nations in international competitions do things differently (De Bosscher *et al*., 2015). In this same study, national sport policy factors such as public expenditure on sport have been found correlated to the international sporting success. It was not the case of other policy factors such as talent identification and development, or athletic (post-)career support.

The first objective of this study was to examine whether clusters of similar elite sport policy systems can be defined among the SPLISS 2.0 nations when using a data-driven multivariate approach. The second objective was to see whether nations in the elite sport policy cluster have similar socio-economic profiles and similar sport performance patterns (level and development).

The current paper enables a better understanding of the pattern of similarities and differences between sport policy systems and examine the association with socio-economic factors (population size and GDP per capita) and the international sporting performance. It may also give new insights to sport policy makers into the group of nations with elite sport policies similar to their own.

## 2 Method

The sample of 16 countries/regions (*nations*) used for this paper consisted of Australia, Brazil, Canada, Denmark, Estonia, Finland, France, Japan, the Netherlands, Northern Ireland, Portugal, South Korea, Spain, Switzerland, as well as Flanders and Wallonia in Belgium.

The scores for the nine SPLISS elite sport policy areas (*pillars*) were retrieved from the SPLISS 2.0 project with reference year 2011 (De Bosscher *et al.*, 2015): financial support (pillar 1), governance, organisation and structure (pillar 2), sports participation (pillar 3), talent identification and development (pillar 4), athletic (post-)career support (pillar 5), training facilities (pillar 6), coach provision and development (pillar 7), (inter)national competition (pillar 8) and scientific research and innovation (pillar 9).

The population count and gross domestic product based on purchasing-power-parity per capita 2011 (GDP per capita) were adopted as socio-economic measures of respectively country size and economic welfare (IMF, 2013; be.stat, 2016; IWEPS, 2015; ONS, 2012; ONS, 2014).

Two measures of international sporting performance were selected for the analysis, for summer and winter sports respectively. First, the sporting performance *level* of a nation was defined by the market share or percentage of the total number of diplomas (top 8 placings) awarded in Olympic Games and world championships (all sports on the Olympic programme) during the period 2009/12, *i.e.* Olympic cycle of four years around the SPLISS reference year. Second, the sporting performance *development* of a nation was defined by the change in market shares between the periods 2009/12 and 2013/16 (percentage points). This measure aimed to assess the impact of the 2011 system on the sporting performance a few years later. Diploma data were retrieved from the Infostrada Podium Performance database. The numbers for Flanders and Wallonia were imputed with two-thirds and one-third of the value for Belgium.

Descriptive statistics provided an overview of the available data. The non-parametric Spearman's correlation coefficient ($r_s$) tested the association between the pillar scores. The sample was partitioned based on the nine sport policy pillar scores. Hierarchical clustering was conducted to get an initial number of sport policy clusters and a first partition (*Ward's* linkage, no standardization). The *k*-means clustering methodology was then applied to settle the final clusters. The ratio "between cluster sum of squares/total sum of squares" ($\eta$) was used as the criterion for comparison among segmentations with a fixed number of clusters (Everitt and Hothorn, 2011; Hennig *et al.*, 2015).

The socio-economic and sporting performance features were compared across the sport policy clusters by using *Kruskal-Wallis* and *Wilcoxon* tests. The analysis was carried out using R (R Core Team, 2017). Two-sided $p$ values of $< .05$ were considered to be significant.

## 3 Results

*Descriptive statistics*

Table 1 presents the variables for analysis. The average pillar scores varied from .47 (pillars 1, finances and 3, participation) to .61 (pillar 5, athletes). The population size, the GDP per capita and the measures of performances varied largely among the nations. Pillars 2 (governance), 5 (athletes), 7 (coaches) and 9 (research) were significantly correlated with one another ($r_s$ from .60 to .81; $p$ lower than .02). In contrast, pillar 8 (competition) was uncorrelated with the other pillars ($p > .05$). In addition, pillars 3 (participation) and 4 (talent) formed a correlated pair ($r_s = .54$, $p = .03$) that was uncorrelated with the other pillars.

**Table 1.** Descriptive statistics of the 16 nations

| | Variable | Mean (SD) | Min | Max |
|---|---|---|---|---|
| Sport policy pillars | 1. Financial support | .47 (.16) | .25 | .70 |
| | 2. Governance, organisation and structure | .48 (.11) | .34 | .69 |
| | 3. Sports participation | .47 (.11) | .33 | .71 |
| | 4. Talent identification and development | .51 (-15) | .18 | .71 |
| | 5. Athletic (post-)career support | .61 (.13) | .34 | .77 |
| | 6. Training facilities | .56 (.13) | .33 | .74 |
| | 7. Coach provision and development | .55 (.14) | .27 | .80 |
| | 8. (Inter)national competition | .55 (.10) | .40 | .78 |
| | 9. Scientific research and innovation | .50 (.18) | .23 | .90 |
| Socio-economic factors | Population [Mio] | 37.6 (53.9) | 1.3 | 197.4 |
| | GDP per capita [USD/1000] | 35.6 (9.6) | 14.3 | 51.6 |
| Sporting Performance* | Perf. level in summer sports [%] | 1.74 (1.42) | 0.17 | 4.71 |
| | Perf. level in winter sports [%] | 2.77 (3.32) | 0 | 11.51 |
| | Perf. dev. in summer sports [% point] | 0.02 (0.25) | -0.53 | 0.39 |
| | Perf. dev. in winter sports [% point] | -0.41 (0.88) | -2.78 | 0.57 |

Note: * missing value for Northern Ireland.

*Sport policy clusters*

Cluster analysis yielded a four-cluster solution. Figure 1 displays the corresponding dendogram ($\eta = 63.5\%$). The first cluster of sport policy systems incorporated Australia (AU), Canada (CA), Japan (JP), Spain (ES), France (FR), and South Korea (KR). CA and JP, as well as KR and FR, had more similar sport policies in this cluster. The second cluster was defined by Switzerland (CH), the Netherlands (NL), Flanders (BE_F), Denmark (DK), and Finland (FI). In this cluster, NL and CH were the first to join together, and DK and FI sport policy systems were also found very similar. The third cluster included Northern Ireland (GB_N), Portugal (PT), Wallonia (BE_W), and Estonia (EE). PT and GB_N were found to be very comparable. EE and BE_W joined them later. The fourth group consisted of Brazil (BR) only.

The upper part of Table 2 displays the sport policy characteristics of the four clusters. The first cluster (AU, CA, JP, ES, FR, KR) was labeled *Leading systems*, because it brought nations together with globally high pillar scores compared with the other clusters, except for pillars 3 (participation) and 4 (talent). The second cluster (CH, NL, BE_F, DK, FI) was labeled *Challenging systems*. It indicated best mean score in pillars 3 (participation) and 4 (talent). The third cluster (GB_N, PT, BE_W, EE) was labeled *Emerging systems*. It was found to have lower mean scores in all but pillars 3 (participation) and 4 (talent). The fourth cluster (BR) was the *Specific system* with a high score in pillar 1 (finance), a good score in pillar 8 (competition) and rather low scores in the remaining pillars.

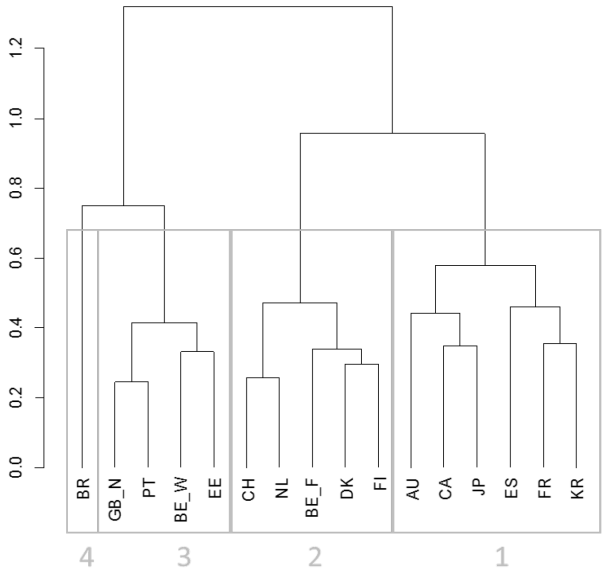


**Figure 1.** Dendogram for the *Ward's* Method. The four clusters of elite sport policy systems are identified with light grey rectangles

*Association with socio-economic factors and sporting performance*

Significant differences were revealed between the sport policy clusters with respect to population size and GDP per capita (Table 2). The nations in the Leading sport policy cluster (AU, CA, JP, ES, FR, KR) showed larger mean population size compared with the Challenging and Emerging clusters. The nations in the Challenging cluster (CH, NL, BE_F, DK, FI) were found to have the largest mean GDP per capita. The nations in the Emerging cluster (GB_N, PT, BE_W, EE) gathered rather small and less wealthy nations. The Specific cluster was made of the very large and less wealthy Brazil.

Different sporting performance features were observed across the sport policy clusters (Table 2 and Figure 2). In summer sports, the nations in the Leading cluster were found to have significantly higher performance level in 2009/12 than nations in the other clusters. However, their market shares decreased (in all nations except for JP) whereas they increased in the Challenging cluster (in all nations except for FI). In winter sports, neither the level nor the evolution differed significantly between the sports policy clusters. However, substantial levels were only observed in the Leading and Challenging clusters. Canada that hosted the Winter Olympics Vancouver 2010 had the highest level and the greatest reduction in performance.

## 4 Discussion

The analysis revealed four sensibly homogeneous and well-separated clusters of similar elite sport policy systems. The elite sport policy clusters have also shown to be associated with population size, GDP per capita and sporting performance in summer sports, but only marginally with the performance in winter sports. The results reinforce the assumption of association between the SPLISS factors and sporting performance in summer sports.

The make-up of the sport policy clusters advocates for three new sets of pillars: (1) the *financial support* pillar 1 (finances), which was the most discriminant pillar distinguishing the clusters; (2) the *fundaments*, made of pillars 3 (participation) and 4 (talent), to develop future potential athletes on the longer term; and (3) the *edifice,* which is formed by the six remaining pillars. These three higher level sets can be a useful contribution to the existing literature on elite sport policies. Together with the correlations between the pillars scores, they suggest that three new composite indicators may contain a large part of the information that differentiates the available sixteen sport policy systems.

**Table 2.** Characteristics of the 4 clusters of elite sport policy systems

| | *M* (*SD*) | | | | Sign. (*p* value) | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | Clust. 1 n=6 (AU, CA, ES, FR, JP, KR) | Clust. 2 n=5 (BE_F, CH, DK, FI, NL) | Clust. 3 n=4 (BE_W, EE, PT, GB_N) | Clust. 4 n=1 (BR) | Overall Clust. 1-3 | Diff 1-2 | Diff 1-3 | Diff 2-3 |
| 1. Financial support | .62 (.06) | .39 (.07) | .29 (.04) | .66 | .003** | .006** | .011* | .049* |
| 2. Governance … | .52 (.10) | .55 (.09) | .36 (.04) | .38 | .022* | .782 | .018* | .014* |
| 3. Sports participation | .42 (.09) | .59 (.10) | .44 (.03) | .35 | .036* | .044* | .669 | .014* |
| 4. Talent identification … | .45 (.12) | .64 (.09) | .52 (.11) | .18 | .053 | .022* | .669 | .086 |
| 5. Athletic (post-)career … | .67 (.08) | .67 (.07) | .50 (.12) | .38 | .045* | .927 | .024* | .037* |
| 6. Training facilities | .67 (.08) | .53 (.09) | .50 (.10) | .33 | .033* | .028* | .032* | .624 |
| 7. Coach provision … | .66 (.09) | .57 (.08) | .44 (.09) | .27 | .017* | .12 | .010* | .082 |
| 8. (Inter)national competition | .61 (.10) | .54 (.10) | .46 (.05) | .57 | .064 | .273 | .025* | .176 |
| 9. Scientific research … | .65 (.18) | .51 (.03) | .32 (.06) | .28 | .010* | .067 | .019* | .014* |
| Population [Mio] | 57.4 (37.2) | 8.4 (4.8) | 4.6 (4.3) | 197.4 | .004** | .006** | .011* | .142 |
| GDP per capita [USD/1000] | 36.7 (4.9) | 44.0 (4.9) | 28.9 (7.4) | 14.3 | .018* | .068 | .088 | .014* |
| Perf. level summer [%] | 3.16 (0.98) | 0.95 (0.68) | 0.30 (0.22) | 1.58 | .012* | .006** | .020* | .101 |
| Perf. level winter [%] | 4.47 (4.11) | 2.87 (2.47) | 0.14 (0.19) | 0 | .082 | .465 | .071 | .053 |
| Perf. dev. summer [%point] | -0.20 (0.21) | 0.17 (0.16) | 0.07 (0.13) | 0.37 | .057 | .017* | .070 | .368 |
| Perf. dev. winter [%point] | -0.76 (1.20) | -0.32 (0.69) | 0.01(0.29) | 0 | .568 | .584 | .302 | .549 |

Note: * $p < .05$ and ** $p < .01$ for overall differences (Kruskal-Wallis test) and differences among pairs (clusters 1-3, Wilcoxon test).

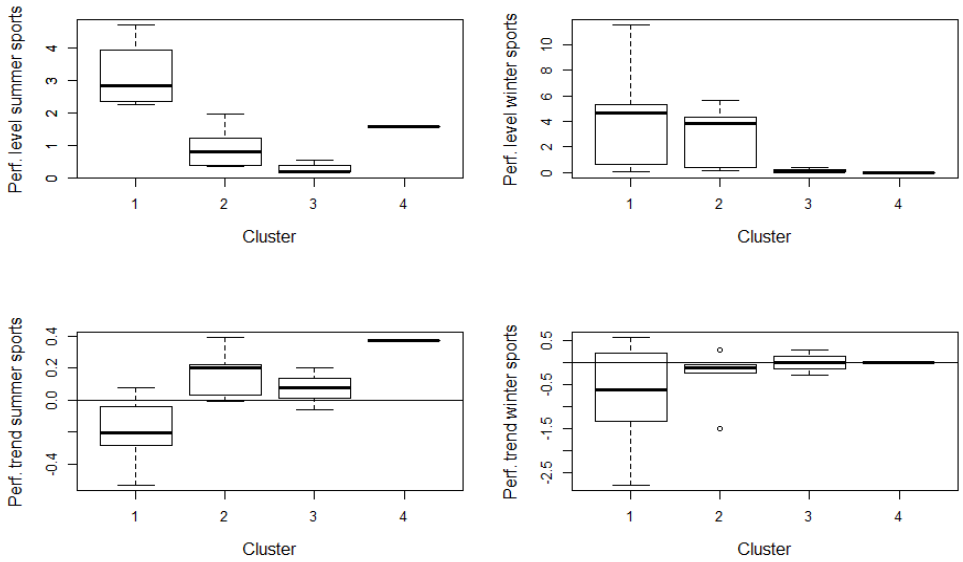


**Figure 2.** Boxplots of the sporting performance level (%) for summer sports (top left) and winter sports (top right), and the performance development (trend, percentage points) for summer sports (bottom left) and winter sports (bottom right) by sport policy clusters

The Leading sport policy systems (high financial input, low/middle fundaments, robust edifice) were the best performing in the period 2009/12 in summer sports. However, these large and moderately wealthy nations may be at a turning point. Their relative performance mostly decreased whereas the nations in the Challenging and Emerging systems - with stronger fundaments - showed the best increasing performances. This suggests that participation and talent identification are key determinants for an increasing performance over time in summer sports.

Some limitations of the study warrant attention. First, the results are only relevant for the available sixteen nations under review and cannot be generalized to other nations. Second, the applied clustering analysis relies on the particular set of the nine SPLISS pillar scores without embedding the whole of the complexity of national elite sport policy systems such as regional and private contributions to elite sport or policies at the sport specific levels. Third, a validation of the clusters was not achievable because not any other set of sport policy scores was available for comparison.

The current study adopts a multivariate perspective to provide new insights in the interplay between sport policy indicators, socio-economic factors and sporting performance. The results offer sport policy stakeholders new opportunities for comparison between nations of the same cluster and with nations of another cluster.

## References


[1] Bernard, A. B. and Busse, M. R. (2004), Who Wins the Olympic Games: Economic Resources and Medal Totals, *The Review of Economics and Statistics, 86* (1), 413-417.

[2] De Bosscher, V., Shibli, S., Westerbeek, H. and van Bottenburg, M. (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.

[3] Everitt, B. and Hothorn, T. (2011). *An introduction to applied multivariate analysis with R*. Springer, New York.

[4] Hennig, C., Meila, M., Murtagh, F. and Rocci, R. (Eds.). (2015). *Handbook of cluster analysis*. Chapman & Hall.

[5] International Monetary Fund [IMF]. (2013), World Economic Outlook Database, retrieved from http://www.imf.org.

[6] Office for National Statistics [ONS] (2012), 2011 Census: Population Estimates for the United Kingdom, March 2011. Retrieved September, 2, 2015, from https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/2011censuspopulationestimatesfortheunitedkingdom/2012-12-17.

[7] Office for National Statistics [ONS] (2014), Regional Gross Value Added (Income Approach): December 2014. Retrieved September, 2, 2015, from https://www.ons.gov.uk/economy/grossvalueaddedgva/bulletins/regionalgrossvalueaddedincomeapproach/2014-12-10.

[8] Reiche, D. (2016). *Success and failure of countries at the Olympic Games*. Routeledge.

[9] Renaud, A., De Bosscher, V. and Kempf, H. (2018). Modelling international sporting performance. The impact of national sport policy factors. Conference paper presented at 50th Annual Conference of the French Society of Statistics, 28 May – 1 June 2018, Paris-Saclay, France.

[10] Statistics Belgium [be.stat] (2016), Population par lieu de résidence, nationalité, état civil, âge et sexe, retrieved from https://bestat.economie.fgov.be/.

[11] Institut wallon de l'évaluation de la prospective et de la statistique [IWEPS](2015), Evolution du PIB en volume, retrieved from http://www.iweps.be/pib-en-volume.

[12] R Core Team (2017), R: A language and environment for statistical computing (Version 3.3.1), Vienna, Austria: R foundation for statistical computing, retrieved from https://www.R-project.org/.