



First Review of IF Governance

2017 ASOIF General Assembly

ASOIF Governance Taskforce
April 2017



First review of IF governance for the General Assembly

Contents

1	Executive summary.....	3
2	Background and objectives	6
3	Methodology	6
4	Headline findings	9
5	Section by section findings.....	11
6	Categorising IFs.....	20
7	Limitations of the study	22
8	Conclusion.....	23
9	Suggested next steps	23
10	Appendices.....	24



1 Executive summary

1.1 Background

In the context of the growing exposure of cases of mismanagement in relation to major sport bodies, the ASOIF General Assembly in April 2016 mandated the Governance Taskforce (GTF) to assist the 28 Summer Olympic International Federations (IFs). The intention was not simply to conduct a compliance exercise but rather to promote a better culture of governance within their structures and help ensure that IFs are fit for purpose, or rapidly achieve that status.

The GTF therefore conducted an evaluation of the current state of governance of the 28 IFs between November 2016 and March 2017 using a self-assessment [questionnaire](#) with independent moderation of the responses. IFs responded very positively to the project with all 28 completing the questionnaire, the large majority doing so ahead of the deadline in January.

1.2 Methodology

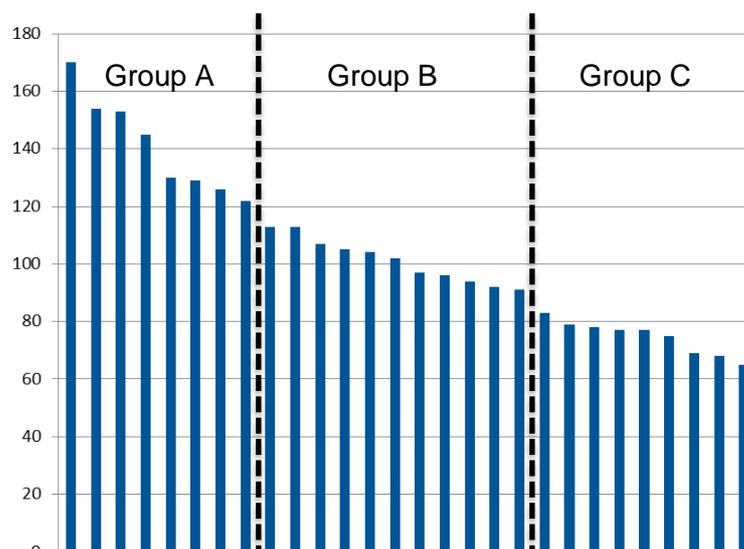
The questionnaire comprised 50 indicators (questions) divided equally among five principles or sections: Transparency, Integrity, Democracy, Development / Solidarity and Control Mechanisms. Each indicator was scored on a scale from 0 (“not fulfilled”) to 4 (“totally fulfilled”). IFs were asked to provide evidence to justify their scores, such as a link to the relevant page/document on their website or a brief explanation.

ASOIF appointed sports governance consultancy I Trust Sport to review the questionnaire responses and moderate scores for consistency. When necessary, scores were adjusted up or down and additional evidence was sought from IF websites. In a few cases, IFs were asked to provide further information.

1.3 Headline findings

There was considerable variation among the 28 IFs with total moderated scores ranging from 65 to 170 out of a theoretical maximum of 200. Three IFs scored over 150 while about half were under 100. The mean score was 104. It should be noted that there are significant differences between IFs in size and scale, which make direct comparisons difficult and to some extent unfair. Several have fewer than 10 full-time staff while a handful have 50 to 120 and one IF is much larger with over 400 employees.

Total moderated scores for IFs in rank order





The IFs can be divided into three groups:

- Group A – 8 IFs scoring 122 to 170
- Group B – 11 IFs scoring 91 to 113
- Group C – 9 IFs scoring 65 to 83

Overall, scores for Transparency were significantly higher than for the other sections. The Integrity and Development / Solidarity sections were the lowest scoring for most IFs. More than one third of the indicators in the Integrity and Development / Solidarity sections across all IFs were scored below 2 on the scale from 0 to 4. By contrast, well over half of the indicators in the Transparency section were scored at 3 or 4.

1.3.1 Transparency

All of the IFs reviewed publish their full Constitution and competition rules effectively, plus information about their members. Only seven IFs published details of allowances and financial benefits. Eighteen IFs published at least one full set of annual, externally audited accounts. Most of the others provided virtually no financial information.

1.3.2 Integrity

This was the most difficult section for IFs, covering several topics which many have only recently started to address, such as policies to counter the threat of match manipulation. Eleven IFs demonstrated no evidence of a confidential reporting mechanism for whistleblowers.

1.3.3 Democracy

All IFs have rules in their Constitution / Statutes regarding the process for elections for Presidents and the majority of the Executive Board. Fifteen IFs had at least some type of term limits in place for the President. Thirteen IFs have an annual Congress or General Assembly, 13 hold a biennial event and in two cases the Congress takes place only every four years.

The indicators on Democracy proved the most challenging part of the questionnaire to assess as IF rules are complex and varied.

1.3.4 Development / Solidarity

IFs scored highest in this section for their provision of education programmes for coaches, judges, referees and athletes. Frequently, there was a lack of published information about the distribution strategy and how funding was targeted. Only 13 IFs were able to provide evidence of formal monitoring or an audit for development funds.

Relatively low scores for Development / Solidarity may partially be explained by the fact that, in some cases, there is a lack of resources to support such activity.

1.3.5 Control Mechanisms

A right of appeal to the Court of Arbitration for Sport was included in the rules of all IFs. There was relatively little evidence of open tendering for contracts offered by IFs, except for event bidding. Sixteen IFs did not meet the criteria of holding “regular open tenders for major commercial and procurement contracts”.



Half of the IFs had an internal audit committee in place with some independent representation. Relatively few IFs could demonstrate specific policies and mechanisms to prevent undue commercial influence on sporting regulations.

1.3.6 Existence of term limits correlated with higher scores

A comparison between the 15 IFs which had some type of term limits in place for the President and the 13 which did not suggests that IFs with term limits had higher average scores. The mean score for those with term limits was 115 as opposed to 92 for those without. It is notable that the introduction of term limits by IFs in recent years has frequently been one component of a wider governance reform programme.

1.4 Limitations of the study

It is believed that the main findings of the study are valid and robust but it is important to acknowledge that the governance assessment has limitations. The questionnaire is just an initial step. Governance is an evolving process, as is governance assessment.

Due to the timetable, there has been little opportunity for dialogue with individual IFs. The questionnaire was standardised and focused on documents and procedures rather than actual behaviour. As such, the study can only provide a partial view – the existence of adequate procedures is an important component of organisational governance but written rules are insufficient on their own.

Many of the limitations could be addressed to some degree in a follow-up study. Tackling other points would require a complementary approach.

1.5 Additional comments

The study revealed large differences between IFs in their governance practices. There were some very impressive examples of high standards but there were also some significant gaps which give cause for concern and need to be addressed.

It will take leadership, commitment and ongoing vigilance for IFs to reach and maintain the high standards of governance which sport deserves.

It is worth emphasising that important reforms are now being adopted rapidly. For example, at least six IFs have introduced term limits since London 2012 and changes approved at Congresses in recent months are in the process of being implemented.

1.6 Suggested next steps

Timing	Activity
After General Assembly	Compile best practice examples
After General Assembly	Identify an official “governance contact” at each IF
Q2-3 2017	Meet individual IFs to review their governance assessments
Sept / Oct 2017	Workshop for designated IF “governance contacts”
By Q4 2017	Revise the questionnaire and assessment process, taking account of feedback
General Assembly 2018	Second iteration of assessment process complete
2018	Explore the creation of a compliance certification system



2 Background and objectives

In the context of the growing exposure of cases of mismanagement in relation to major sport bodies, the ASOIF General Assembly in April 2016 mandated the Governance Taskforce (GTF) to assist the 28 Summer Olympic International Federations (IFs). The intention was not simply to conduct a compliance exercise but rather to promote a better culture of governance within their structures and help ensure that IFs are fit for purpose, or rapidly achieve that status.

The GTF therefore conducted an evaluation of the current state of governance of the 28 IFs between November 2016 and March 2017 using a self-assessment [questionnaire](#) with independent moderation of the responses. IFs responded very positively to the project with all 28 completing the questionnaire, the large majority doing so ahead of the deadline in January.

As [explained](#) at the start of the project, the questionnaire consisted of 50 fairly simple and measurable indicators covering five principles or sections: Transparency, Integrity, Democracy, Development/Solidarity and Control Mechanisms. There was also an initial section on Guiding Codes, referencing the foundation documents of the Olympic Movement, such as the Olympic Charter and the World Anti-Doping Code.

3 Methodology

3.1 Scoring

Each of the 50 indicators in the questionnaire incorporated a separate definition for scores on a scale from 0 to 4. The scores in each case were designed to assess the level of fulfilment of the indicator by the IF, as follows:

- 0 – Not fulfilled at all
- 1 – Partially fulfilled
- 2 – Fulfilled
- 3 - Well-fulfilled according to published rules/procedures
- 4 – Totally fulfilled in a state of the art way

IFs were asked to provide evidence to justify their scores, such as a link to the relevant page/document on their website or a brief explanation.

The initial Guiding Codes section was slightly different in format but a similar process applied.

3.2 Independent moderation

ASOIF appointed sports governance consultancy I Trust Sport to support the project. I Trust Sport's task was to review the questionnaire responses; to moderate the scores to ensure as much consistency as possible; and to produce analysis for this report.

Scores were checked against the defined criteria in the questionnaire for each indicator for all 28 responses. Evidence provided by IFs was also checked (such as references to clauses in the Constitution or specific web pages) and, where evidence was absent or incomplete, additional information was researched from IF websites.



When necessary, scores were adjusted up or down to reflect the independent assessment of the moderator, based on the evidence available. The aim was to be consistent and fair.

The questionnaire is new and so it is not surprising that there was variation in the interpretation of some of the indicators by different IFs, resulting in a need to adjust scores.

During the course of the moderation process, a number of policy decisions were applied regarding the scoring of specific indicators (with the approval of the GTF) to provide added consistency (see below and 10.2 in the appendix).

It is important to note that the assessment represents a snapshot in time. Questionnaires were returned to ASOIF by IFs in late January and early February 2017. The moderation process ran from 27 January to 9 March, during which time documents were downloaded and pages of IF websites reviewed. As is to be expected at this point in the four year cycle, many IFs were in the process of implementing governance reforms or preparing for Congresses. The analysis is based on regulations that were in place on the day of moderation, not taking account of future changes, even where these were imminent and/or certain to be implemented. This seemed to be the fairest approach.

The findings in paragraphs 4 to 6 below relate to the five sections Transparency, Integrity, Democracy, Development / Solidarity and Control Mechanisms. The initial section on Guiding Codes is effectively covered elsewhere in the questionnaire and so it has not been analysed separately.

3.3 Outcomes of moderation

Change in scores after moderation:

	Self-assessed score	Moderated score
Mean for total*	122	104
Median for total*	122.5	99.5
Mean per indicator	2.44	2.08
Median per indicator	2.45	1.99

Maximum increase	+23 (moderated score is 23 above self-assessed score)
Maximum decrease	-81 (moderated score is 81 below self-assessed score)
Mean change	-18
Median change	-12.5
Median change %	-13%

(*) Note on mean and median:

The mean is the sum of the figures divided by the number of figures (so divided by 28 to calculate a mean score for each IF). The median is the mid-point when a set of numbers are listed from smallest to largest (so half-way between the 14th and 15th if 28 IF scores are being considered). The median is less impacted by an unusually high or low number in the series. Both mean and median are used in this report.

The moderated scores of all but two of the IFs were lower than the self-assessed scores. As the questionnaire is new and some of the indicators could be interpreted in different ways, it is understandable that there were inconsistencies in the responses, which the moderation process attempted to address. The fact that quite a number of scores were moderated down should not be interpreted as a criticism of the work of the IFs in completing the questionnaire. As ASOIF acknowledges, the questionnaire lacked clarity in places and can be improved significantly, building on the experience gained and feedback received (paragraph 9 below proposes next steps).



In the cases where there were very large differences between the moderated and self-assessed scores (six IFs were marked down by 30 or more points), the discrepancy tended to relate to the interpretation of “publishing”: several IFs allocated scores of 4 for indicators relating to the publication of information such as Congress minutes when details were distributed to members only. In the moderation process high scores were awarded for these indicators only when information was published on the IF website.

The decision to base assessment on regulations that were in place on the day of the review (see 3.2 above) resulted in a significant number of scores being moderated down because several IFs understandably wanted to take into account governance reforms that were due to be implemented in the coming weeks or months. If another assessment exercise takes place in future, improved governance will naturally be acknowledged.

3.4 Allowing a margin of error

The scoring system gave the analysis a degree of objectivity. However, in many cases there was room for debate.

For example, when moderating one questionnaire, the scores for a total of nine out of 50 questions were initially marked as being uncertain (in each case the doubt concerned only a single point). In every case a decision was made and the total score was marked down by five points but another analyst could have come to a slightly different conclusion.

The median reduction in score across the 28 questionnaires was -12.5 points. On the basis that some of these judgements could be debatable, each IF total score should be understood to have a margin of error from -7 to +7.

For more details of the moderation process and the policies applied, see 10.2 in the appendix. This includes a statistical check to test the level of agreement between the self-assessed and moderated scoring – what is known as “inter-coder reliability”. The test concluded that inter-coder reliability was at a satisfactory level.

3.5 Note on varying size and scale of IFs

The 28 IFs which are members of ASOIF are very varied in their size and the scale of their operations. Many are rather small – about half have fewer than 20 full-time staff and several have fewer than 10. An estimated seven IFs have 50 to 120 or so members of staff and one is much larger, employing over 400 people. (Staff numbers are taken from an unpublished study in 2015 and are approximate.)

The scale of activity also covers quite a range. While some IFs manage and run global competition series with substantial prize money, others do not own the rights to the largest events in their sports.

Due to this variation, caution is needed when attempting to make direct comparisons between IFs.

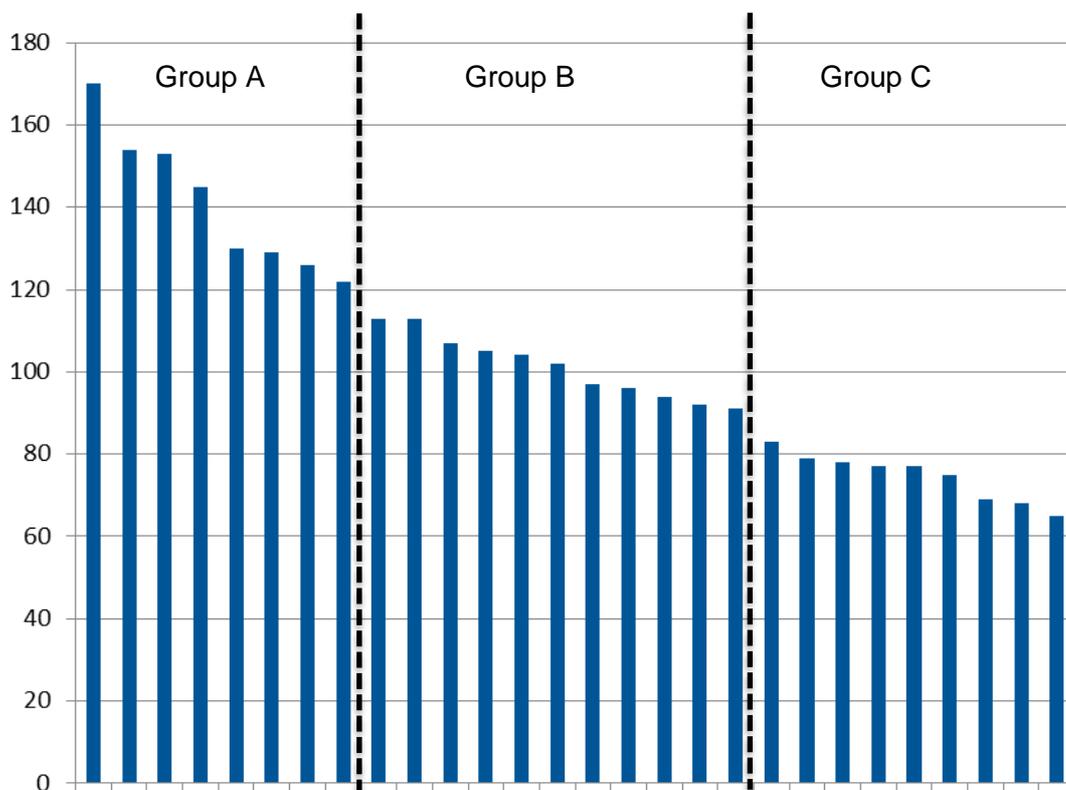
4 **Headline findings**

The analysis of the questionnaire data has been anonymised so that scores from individual IFs are not readily identifiable. While the chosen approach does limit the potential for external scrutiny, it is believed to be fair considering the limitations of the project (paragraph 7 below) and this study is intended only as the first step in a committed attempt to improve the governance of IFs.

Please note that all of the analysis is based on moderated scores, not self-assessed scores.

4.1 **Overall moderated scores**

Moderated total scores for each IF in rank order



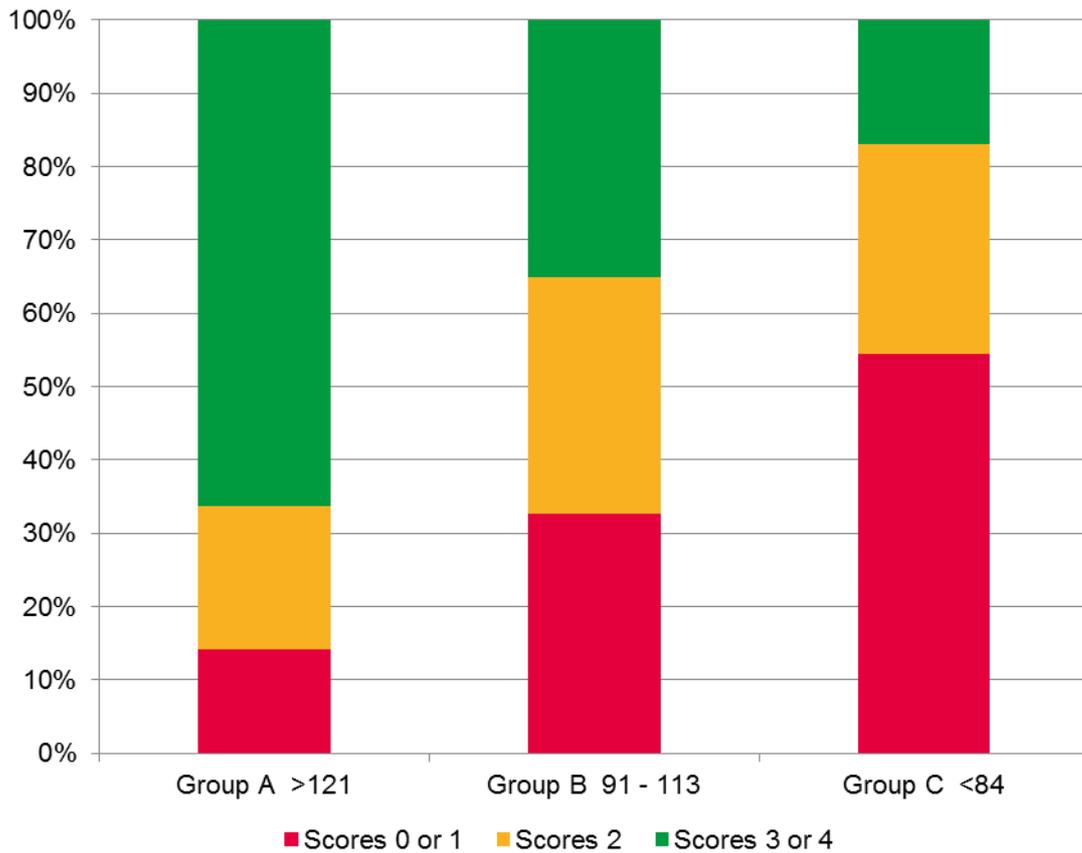
There was considerable variation among the 28 IFs. Total moderated scores ranged from 65 to 170 out of a theoretical maximum of 200. Three IFs scored over 150, which is an average of more than 3 out of 4 for each indicator. Half scored under 100, below an average of 2 per indicator. However, allowing for a margin of error from -7 to +7 in each total score, three more IFs could reach this threshold.

Based on the scores, it seems realistic to divide the IFs into three groups. The strongest eight (Group A) each scored at least 122. Considering the margin of error, the order of the leading group could vary but probably not the composition. In the leading group, half or more of the 50 indicators tended to be scored at 3 or 4 (“well-fulfilled” or “totally fulfilled”).

As might be expected, in the middle ranking group (Group B) the differences are fairly small. There were 11 IFs with scores ranging from 91 to 113. Averaging a score of close to 2 across the 50 indicators, these IFs mostly had a fairly even spread of indicators from 0 to 4.

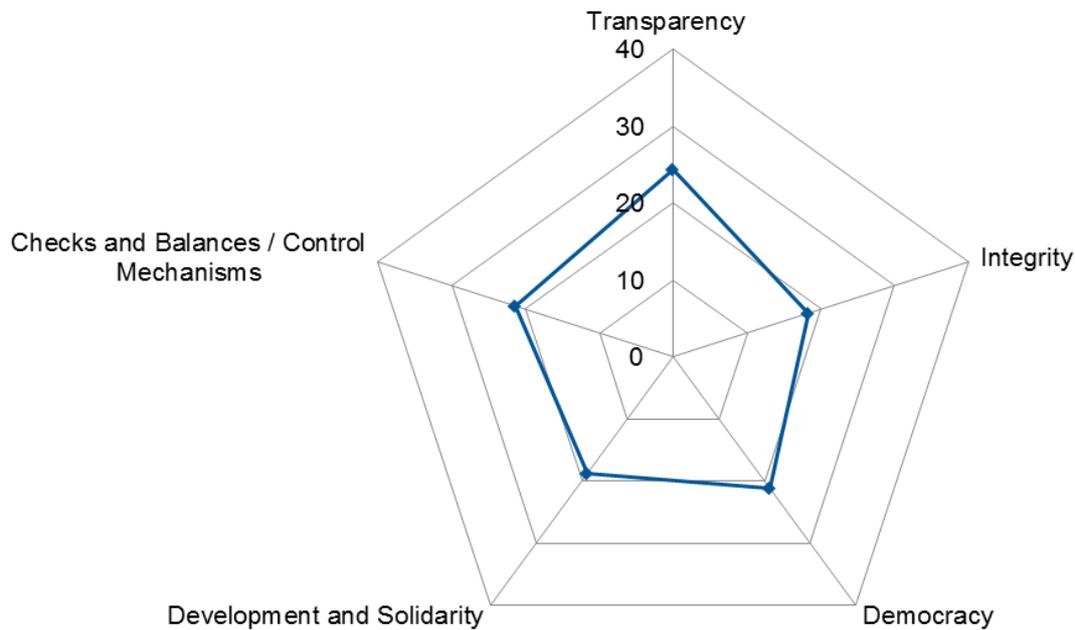
The weakest group (Group C) consists of nine IFs which scored no more than 83. Even allowing for the margin of error, it is unlikely that any of these could justifiably be included in Group B. At this level, 20 or more of the 50 indicators were generally scored at 0 or 1 (“not fulfilled at all” or “partially fulfilled”).

Percentage distribution of scores for Groups A, B and C



5 Section by section findings

Median scores by section (theoretical maximum of 40)



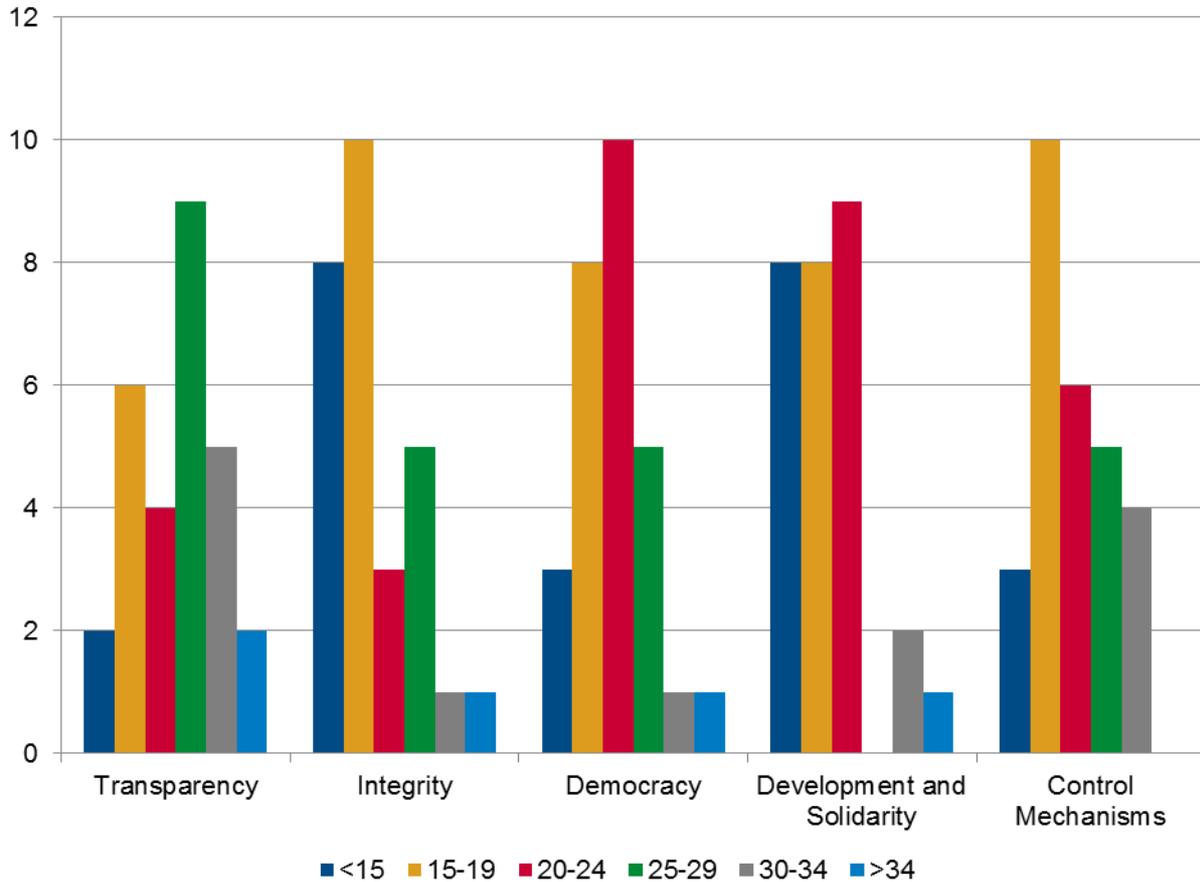
Section	Min	Max	Mean	Median
Transparency	14	39	24	25
Integrity	6	37	18	16
Democracy	13	38	21	21
Development / Solidarity	7	38	19	17.5
Checks and Balances / Control Mechanisms	12	34	21	20.5

Overall, scores for Transparency were significantly higher than for the other sections with a median of 25. Sixteen IFs had their best score for Transparency. The Integrity section was the lowest scoring or joint lowest scoring for 11 out of 28 IFs, the median score being 16. Development / Solidarity was lowest or joint lowest for 10 IFs. The differences between the spread of scores for Democracy and Checks and Balances / Control Mechanisms were fairly small. The Checks and Balances / Control Mechanisms section had the lowest maximum score with 34.



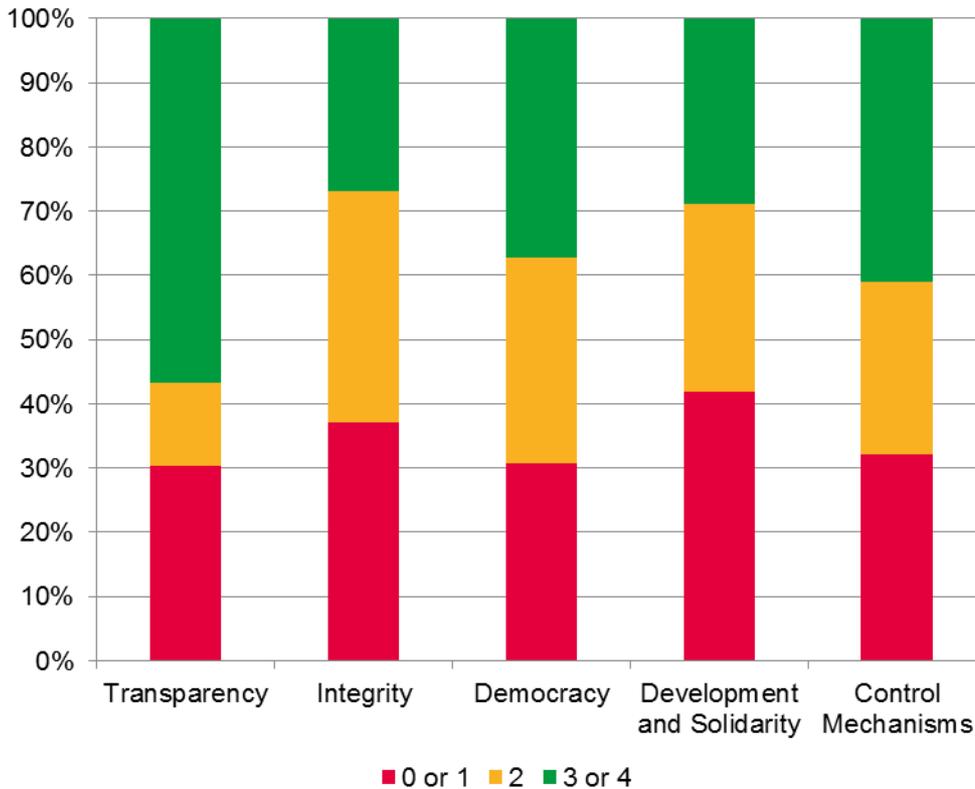
Distribution of scores by section

Count of IFs



Integrity and Development / Solidarity accounted for 16 of the 24 section scores below 15. Only three IFs recorded more than 24 for Development / Solidarity. One hypothesis is that achieving a high score in this section is at least partially dependent on having sufficient resources available. In the Democracy section, scores were clustered in the middle with 23 of 28 IFs registering between 15 and 29. There were only five scores of over 34 for a specific section among the 28 IFs (out of 140 in total).

Distribution of scores for individual indicators by section (%)



Section	0 or 1	2	3 or 4
Transparency	30%	13%	57%
Integrity	37%	36%	27%
Democracy	31%	32%	37%
Development / Solidarity	42%	29%	29%
Control Mechanisms	32%	27%	41%
Total	34%	27%	38%

Over one third of all of the indicators in the Integrity and Development / Solidarity sections across all IFs were scored below 2 (10 indicators for each of 28 IFs so the total is 280 indicators per section). The Transparency section recorded the smallest proportion of indicators scored below 2 at 30%. Well over half of the indicators in the Transparency section were scored at 3 or 4.

Across the whole study, about a third of all indicators were scored below 2. However, 38% achieved a score of 3 or 4, which signifies that there were more instances of governance requirements being well or totally fulfilled than fulfilled partially or not at all.



5.1 Transparency section

Mean scores by indicator

Indicator	Topic	Mean
2.1	Statutes, rules and regulations	3.36 (highest)
2.2	Organisational charts for staff, elected officials and committee structures, and other relevant decision-making groups including Remuneration Committee	2.89
2.3	Vision, mission, values and strategic objectives	2.18
2.4	A list of all national member federations with basic information for each	3.25
2.5	Details of elected officials with biographical info	2.25
2.6	Annual activity report and main events reports	2.32
2.7	Annual financial reports following external audit	2.07
2.8	Allowances and financial benefits of elected officials and senior executives	1.11 (lowest)
2.9	General Assembly agenda with relevant documents (before) and minutes (after) with procedure for members to add items to agenda	2.46
2.10	A summary of reports/decisions taken during Executive Board and Commission meetings and all other important decisions of IF	2.50

Transparency was the strongest section of the five in the questionnaire.

All of the IFs reviewed published their full Constitution and competition rules effectively, plus information about their members. 2.1 was the highest scoring of the 50 indicators in the study.

By contrast, relatively few IFs published details of allowances and financial benefits for elected officials and senior executives. Only seven IFs published allowances/per diems/benefits for officials or senior staff (scoring 2 or more for indicator 2.8).

Eighteen out of 28 IFs published at least one set of annual, externally audited accounts (scoring 2 or more for indicator 2.7). Most of the others provided virtually no financial information on their websites.

5.2 Integrity section

Mean scores by indicator

Indicator	Topic	Mean
3.1	Has a unit or officer in charge of ensuring that the IF abides by the IOC Code of Ethics and/or the IF's own Code of Ethics	2.21
3.2	Has a unit or officer in charge of ensuring that the IF abides by the WADA World Anti-Doping Code	2.68 (highest)
3.3	Complies with the Olympic Movement Code on the Prevention of the Manipulation of Competitions	2.25
3.4	Has a programme or policies designed at ensuring that the IF member associations function in accordance with all recognised ethical codes and principles	1.75
3.5	Establish confidential reporting mechanisms for "whistle blowers" with protection scheme for individuals coming forward	1.07
3.6	Respect principles of sustainable development and regard for the environment	1.68
3.7	Put in place integrity awareness/education programmes	1.46
3.8	Provide for appropriate investigation of threats to sport integrity	2.29
3.9	Cooperate with relevant public authorities (such as Interpol, Council of Europe, UN Office on Drugs and Crime) on integrity matters	0.93 (lowest)
3.10	Make public all decisions of disciplinary bodies and related sanctions, as well as pending cases where applicable	1.96

The indicator on anti-doping activity produced the highest mean score in the Integrity section.

The section included several indicators on topics which have only recently started to be addressed by IFs, such as whistleblower mechanisms (indicator 3.5) and policies to counter the threat of match manipulation (3.3, 3.7, 3.8 and 3.9). It is therefore perhaps not surprising that this section was the lowest scoring for 11 of 28 IFs.

Eleven IFs scored 0 for indicator 3.5, implying that there was no evidence of a confidential reporting mechanism for whistleblowers.

In their response to indicator 3.3, most IFs referenced their co-operation with the International Olympic Committee's Integrity Betting Intelligence System (IBIS).

Only two IFs scored 30 or more for the section and eight scored under 13. 3.9 was the lowest scoring of the 50 indicators in the questionnaire.



5.3 Democracy section

Mean scores by indicator

Indicator	Topic	Mean
4.1	Election of the President and a majority of members of all executive bodies	2.89
4.2	Clear policies to ensure election candidates can campaign on balanced footing including opportunity for candidates to present their vision/programmes	2.07
4.3	Election process with secret ballot under a clear procedure/regulation	3.11 (highest)
4.4	Make public all open positions for elections and appointments including the process for candidates and full details of the roles, job descriptions, application deadlines and assessment	1.75
4.5	Establishment and publication of eligibility rules for candidates for election together with due diligence assessment	1.86
4.6	Term limits for elected officials*	1.00 (lowest)
4.7	Provide for the representation of key stakeholders (e.g. active athletes) in governing bodies. Due regard shall be paid to gender representation and the enactment of policies encouraging gender equality	2.32
4.8	Main decisions are taken on basis of written reports supported by criteria with opportunity for secret ballot at request of voting constituents	2.07
4.9	Defined conflict of interest policy with exclusion of members with a manifest, declared or perceived conflict	2.11
4.10	Governing bodies meet regularly	2.00

All IFs have rules in their Constitution or Statutes regarding the election process. In most cases, the President is elected by all of the members of the IF, as are the majority of the Executive Board or equivalent.

Thirteen out of 28 IFs had no term limits in place, scoring 0 for indicator 4.6. It was the second lowest scoring of all 50 indicators (see also 6.2 below).

Thirteen IFs have an annual Congress or General Assembly and so scored 2 or more for indicator 4.10. Most of the rest hold a biennial event. In two cases the Congress takes place every four years. There have been a significant number of Extraordinary Congresses in recent months.

The indicators on Democracy proved the most challenging part of the questionnaire to moderate. Three indicators in particular required the application of specific policies to ensure consistency (4.4, 4.5 and 4.7 - see 10.2 below). It is difficult to do justice to the level of complexity and variation in election regulations among IFs in only a handful of indicators (4.1, 4.2, 4.3, 4.5, 4.6 and 4.7).



Some of the variation in election processes is explained and justified by the history of the IFs (for example, the relationship between the IF and professional leagues or player associations in the sport).

The general subject of democratic processes in IFs might merit a dedicated piece of research.

(*) Note on “term limits”: Term limits refer to a restriction on the length of time an official such as a President or Executive Board member can serve in an organisation. Term limits may specify a period of years or the number of times an individual can be eligible for re-election. In order for indicator 4.6 to be fulfilled (scoring 2 or above), there had to be a limit at least as strict as permitting a maximum of three terms of four years for an individual in one role.



5.4 Development and Solidarity

Mean scores by indicator

Indicator	Topic	Mean
5.1	Transparent process to determine allocation of resources in declared non-profit objectives	1.86
5.2	Redistribution policy and programmes for main stakeholders	2.14
5.3	Monitoring / audit process of the use of distributed funds	1.57
5.4	Existence of environmental responsibility policy and measures	1.71
5.5	Existence of social responsibility policy and programmes	2.00
5.6	Education programmes and assistance to coaches, judges, referees and athletes	3.00 (highest)
5.7	Solidarity programmes pay due regard to gender and geographical representation through internal guidelines	1.61
5.8	Legacy programmes to assist communities in which events are hosted	1.64
5.9	Anti-discrimination policies on racial, religious or sexual orientation	1.75
5.10	Cooperation with relevant public authorities (such as UN Environment Programme, ICRC, World Health Organisation, UN High Commissioner for Refugees) on social responsibility issues	1.50 (lowest)

IFs scored highest for their provision of education programmes for coaches, judges, referees and athletes, which seemed to be well-established in most cases.

Solidarity programmes were generally in place but frequently there was a lack of published information about the distribution strategy, how funding was targeted and the extent of any monitoring. For example, 15 out of 28 IFs scored 0 or 1 for indicator 5.3, meaning that funding recipients may be required to provide a report but there was no evidence of formal monitoring or an audit.

Indicator 5.10 had the lowest mean score in this section. Relatively few IFs demonstrated co-operation with public authorities. However, the indicator did not reference other bodies with which IFs may co-operate in complementary development activity such as National Olympic Committees through the network of the member National Federations.

Development / Solidarity was the lowest or joint lowest scoring section for 10 IFs, which may be explained by the fact that activity is limited in some cases due to availability and allocation of resources.



5.5 Control Mechanisms section

Mean scores by indicator

Indicator	Topic	Mean
6.1	Establish an internal ethics committee with independent representation	2.29
6.2	Establish an audit committee that is independent from the decision-making body and reports to members directly	1.64
6.3	Adopt accounting control mechanisms and external financial audit	2.29
6.4	Adopt policies and processes for internal control and risk management	2.54
6.5	Adopt policies and mechanisms to prevent commercial interests from overriding sporting regulations e.g. conduct of draws	1.71
6.6	Observe open tenders for major commercial and procurement contracts	1.36 (lowest)
6.7	Decisions can be challenged through internal appeal mechanisms on the basis of clear rules	2.18
6.8	Due diligence and effective risk management in bidding requirements, presentation, assessment and allocation of main events	2.00
6.9	Awarding of main events follows an open and transparent process	2.36
6.10	Internal decisions can be appealed with final recourse to the Court of Arbitration for Sport	3.07 (highest)

The right of appeal to the Court of Arbitration for Sport is included in the rules of all IFs which were reviewed. Most but not all sports have had experience of at least one case.

There is relatively little open tendering for contracts offered by IFs, except for event bidding. Sixteen of the 28 IFs scored 0 or 1 for indicator 6.6, meaning that they did not meet the criteria of holding “regular open tenders for major commercial and procurement contracts”.

The concept of an audit committee is fairly new for many IFs. Half of them had an internal audit committee in place with some independent representation (not staff or members of the executive), scoring 2 or more for indicator 6.1.

Relatively few IFs could demonstrate having specific policies and mechanisms in place to prevent undue commercial influence on sporting regulations, although there was often an indirect reference in the Code of Ethics.



5.6 Wider impact of Transparency

While only one of the five sections in the questionnaire was specifically dedicated to Transparency, the extent to which an IF publishes information had a wider impact on scores. For example, when an annual report was published, it generally included details of development activity (relevant to the Development / Solidarity section). Similarly, audited accounts often provided an outline of internal controls and risk management (covered in the Control Mechanisms section). For a number of the IFs with scores at the lower end of the scale, a shift to publishing more information could provide a substantial boost without the need for constitutional amendments.

6 Categorising IFs

6.1 Categorising IFs by resources – limited evidence of correlation with scores

As referenced in 3.5 above, there is considerable diversity among IFs in their budgets, the number of professional staff employed and in the scope of their operations. It is plausible that these differences could have an impact on governance structures and procedures. It might also be reasonable to have different expectations which are proportionate to the resources available and the scope of activity.

Analysis was attempted to look for correlations or patterns in scores among groups of IFs with larger or smaller salary budgets using information from audited accounts.

Based on the information available, no very strong trends could be identified. There are examples of IFs operating with a small number of staff which scored very well and other cases of IFs with access to considerable resources which were at the lower end of the scale.

In the Integrity and Control Mechanisms sections there was some evidence that IFs with a salaries budget of under 2.5m CHF (about 16 IFs) scored slightly lower. This may be because some of the IFs with limited resources have fewer policies and procedures in place to tackle recently identified priorities such as the prevention of the manipulation of sports competitions.

Another idea was to divide IFs into two groups based on staffing numbers taken from an unpublished study in 2015. The groups consisted of 20 IFs with fewer than 30 staff and seven with over 30 staff (one outlier with a much larger number of staff than the others was excluded from the analysis). The IFs that had fewer than 30 staff scored significantly lower on average. However, there were a number of exceptions and the staffing numbers could not be treated as definitive. For example, the method of counting remotely based workers and consultants is not easily compared between IFs.

Using the data collected, it is not clear that there was a straightforward or fair method to divide IFs into categories based on their resources which would allow an additional level of analysis. However, a future study might consider methods of doing so.

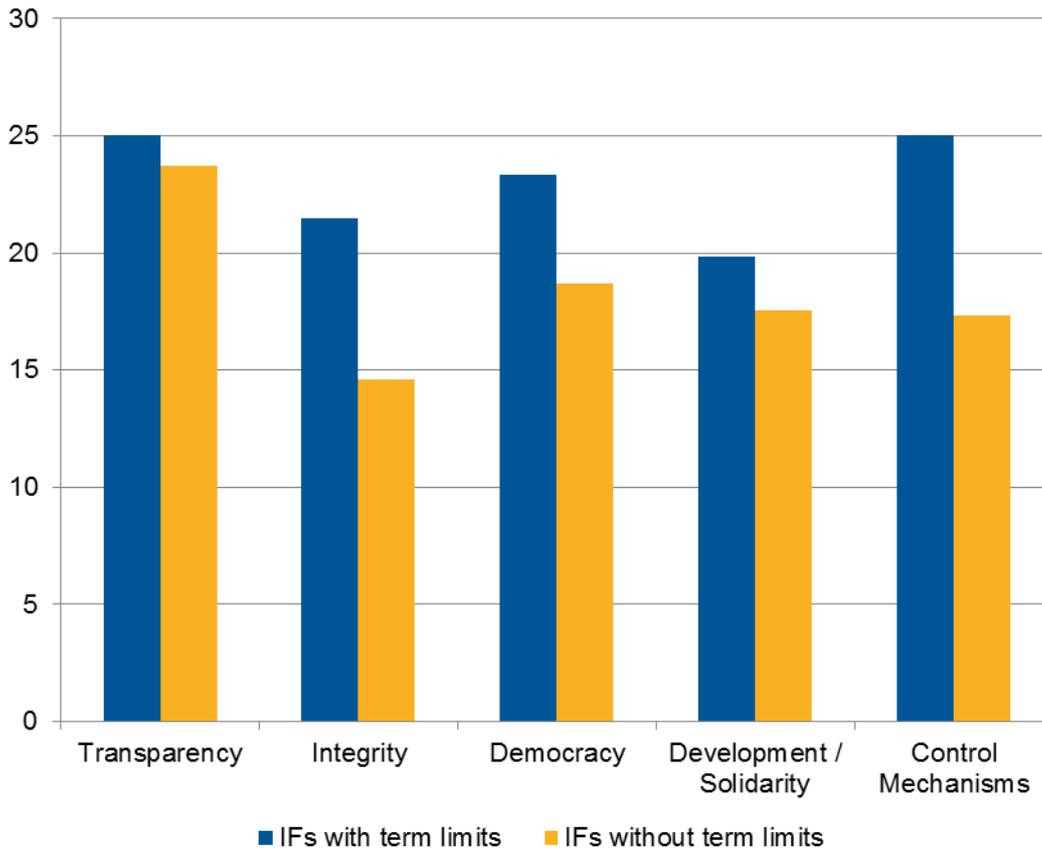
6.2 Existence of term limits – evidence of correlation with scores

Fifteen out of 28 IFs had at least some kind of term limit in place for the President, although precise rules vary considerably (see also the note on term limits in 5.3 above). In many cases, the term limits have been introduced within the last few years as one element of a governance reform programme, which may have had an impact on scores across the questionnaire.

A comparison of IFs with no term limits (which scored 0 for indicator 4.6) to those with some type of limit in place suggests significant differences:

1.3.3

Mean scores by section for IFs with and without term limits in place



Term limits?	No. of IFs	Mean scores					
		Total	Transparency	Integrity	Democracy*	Development / Solidarity	Control Mechanisms
Some term limit(s)	15	115	25	21	23	20	25
No term limits	13	92	24	15	19	18	17

(*) The Democracy score obviously includes 0 for indicator 4.6 in the case of IFs with no term limits and more than 0 for IFs that do have term limits, which explains part of the difference.

On average, an IF with some type of term limits in place reaches the higher end of Group B (at around 115 points). By contrast, IFs without term limits are at the lowest end of Group B (about 92 points). The mean scores for each separate section are also higher among the IFs



which have some type of term limits in place, most notably for Integrity and Control Mechanisms.

7 Limitations of the study

It is believed that the main findings of the study are valid and robust but it is important to acknowledge that this is the first governance assessment project undertaken by ASOIF and there are limitations.

The self-assessment questionnaire should be regarded as an initial step and a snapshot in time. Governance is an evolving process, as is governance assessment. Concepts of best practice and public expectations of governance change over time. As was evident in the study, many IFs are currently in the process of making significant changes.

Due to the timetable, there has been little opportunity for dialogue with individual IFs about the questionnaires before the ASOIF General Assembly. Consequently, there may be errors in the assessment due to information that has not been provided or has been misinterpreted.

Desk analysis of documents, procedures and structures does not take account of actual behaviour and organisational culture. For example, a process may be documented but not implemented in practice, or written materials which are updated infrequently may not yet reflect an important change that has been agreed.

The questionnaire was deliberately limited to 50 indicators to make the task of completion manageable. Inevitably, it did not address every relevant issue. Some important and high profile topics which were not covered in detail include gender equity, evidence of criminal activity, and welfare issues including safeguarding children and vulnerable adults.

Most of the scoring definitions described a predominantly “quantitative” rather than “qualitative” view – for example the existence of rules for an internal appeals process rather than how effectively the rules work (6.7).

Scores were not weighted. It is possible for good scores achieved across various topics by an IF to mask a serious failure in one specific area, or for gaps in non-core areas to bring down an overall score although the basics were covered well.

The thematic division between sections in the questionnaire was pragmatic but is not claimed to be scientific.

In places, the questionnaire lacked clarity and there were differences in the interpretations of a few indicators by the respondents. This was dealt with as far as possible in the moderation process.

As explained elsewhere, direct comparisons between IFs could be unfair in some cases. While every effort was made to be consistent, it is unrealistic to claim that a score of 3 for a particular IF on one indicator is exactly the same as a 3 for another. There is a degree of subjectivity in the scoring, which explains the need to accept a margin of error.

The questionnaire was Olympic-focused and did not consider Paralympic issues. Eleven of the 28 IFs are responsible for Paralympic sport so it would require further work to adopt a consistent approach in this area.

Many of these limitations could be addressed to some degree in a follow-up study. Tackling others would require a complementary approach.



8 Conclusion

The ASOIF Governance Taskforce believes that the governance assessment project has been a very worthwhile exercise so far and looks forward to continuing the work. The questionnaire was positively received by the IFs, who responded for the most part on time and in sufficient detail to enable the moderating process and analysis to progress smoothly. Improvements can be made to the questionnaire but the current version provided a considerable quantity of useful data to work with.

The study revealed large differences between IFs in their governance practices. There were some very impressive examples of high standards but there were also some significant gaps which give cause for concern and need to be addressed.

It will take leadership, commitment and ongoing vigilance for IFs to reach and maintain the high standards of governance which sport deserves. Policies and processes are important but they are not enough in themselves – culture and behaviour at all levels in an organisation play a huge role.

On a positive note, it is worth emphasising that important reforms, which were perhaps overdue, are now being adopted rapidly. One example is the introduction of term limits for elected officials.

As explained in 5.3 above, in the current study 15 out of 28 summer IFs had some type of term limits in place. In at least six cases, the term limit rules have been introduced since London 2012.

Another sign of progress is evident from the fact that a number of IFs have published audited accounts for the first time in the last couple of years.

Several IFs explained in their questionnaires that they were in the process of implementing reforms recently approved at post-Olympic Congresses. Others drew attention to proposed measures to be voted on in the course of 2017. It is encouraging to see that the progress looks set to continue.

9 Suggested next steps

The ASOIF Governance Taskforce plans to continue with the governance assessment project. The intended next steps are as follows:

Timing	Activity
After General Assembly	Compile best practice examples across a range of aspects of governance, drawn from the findings of the study
After General Assembly	Identify an official “governance contact” at each IF
Q2-3 2017	Meet individual IFs to review their governance assessments
Sept / Oct 2017	Workshop for designated IF “governance contacts”
By Q4 2017	Revise the questionnaire and assessment process, taking account of feedback
General Assembly 2018	Second iteration of assessment process complete
2018	Explore the creation of a compliance certification system



10 Appendices

10.1 International Federations

International Boxing Association (AIBA)
Badminton World Federation (BWF)
Fédération Equestre Internationale (FEI)
Fédération Internationale d'Escrime (FIE)
Fédération Internationale de Basketball (FIBA)
Fédération Internationale de Football Association (FIFA)
Fédération Internationale de Gymnastique (FIG)
Fédération Internationale de Hockey (FIH)
Fédération Internationale de Natation (FINA)
Fédération Internationale de Volleyball (FIVB)
Fédération Internationale des Sociétés d'Aviron (FISA)
International Association of Athletics Federations (IAAF)
International Canoe Federation (ICF)
International Golf Federation (IGF)
International Handball Federation (IHF)
International Judo Federation (IJF)
International Shooting Sport Federation (ISSF)
International Table Tennis Federation (ITTF)
International Tennis Federation (ITF)
International Triathlon Union (ITU)
International Weightlifting Federation (IWF)
Union Cycliste Internationale (UCI)
Union Internationale de Pentathlon Moderne (UIPM)
United World Wrestling (UWW)
World Archery (WA)
World Rugby (WR)
World Sailing (WS)
World Taekwondo Federation (WTF)



10.2 Further explanation of moderation process

10.2.1 Assumptions made in conducting moderation and calculating scores

- The reviews were based only on responses provided in the questionnaire, material on the relevant IF website and (in a small number of cases) on supplementary documents submitted by IFs along with the questionnaire
- Scores were based on sections 2-6 of the questionnaire, excluding section 1 (Guiding Codes). Indicators 3.1, 3.2 and 3.3 effectively served as a proxy for 1.2, 1.4, 1.5, 1.7 and 1.8
- Moderated scores were based on regulations which were in place on the day on which the questionnaire was reviewed - credit was not given for planned future reforms. This had a negative impact on some scores but seemed the fairest approach
- The assessment took some account of what seemed proportionate to the resources of the IF (e.g. in terms of the approach to development programmes) but a modest size/budget should not excuse poor practice

10.2.2 Indicative example of moderating scores

Below an anonymised example is provided of the moderation process for a specific indicator using the self-assessed and moderated scores for three separate IFs.

Indicator 2.7 - Annual financial reports following external audit

Mean score: 2.07

Scores	Definition in questionnaire distributed
0	No
1	Some financial information published on IF website
2	Publication of externally audited financial reports on IF website
3	Publication of audited financial reports, easy to find on IF website
4	Publication of state of art audited financial reports, easy to find on IF website, extra data, management letter

Example IF A

Self-assessed score	Evidence in questionnaire response
4	IF states that annual accounts are provided to members in Congress papers
Moderated score	Rationale for moderated score
0	No financial information published on website

Example IF B

Self-assessed score	Evidence in questionnaire
3	Hyperlink provided to annual accounts which are among Congress papers in the dedicated section on the website
Moderated score	Rationale
2	Files are difficult to find as it is necessary to navigate through several pages. There is very little explanatory information beyond the accounts themselves

Example IF C

Self-assessed score	Evidence in questionnaire
4	Hyperlink provided to download a detailed financial report and audited accounts; there are documents for several previous years
Moderated score	Rationale
4	Meets criteria; accounts are relatively easy to find in a logical process from the home page

10.2.3 Overlapping indicators

There were three pairs of indicators which overlapped fairly closely in subject matter. While it was possible to differentiate between them, the distinctions were subtle and there was some inconsistency in the way they were answered. A significant number of IFs provided the same answer to both questions.

For each of the pairs of indicators below, in the case of doubt, IFs were awarded the same score for both indicators.

Pairs of overlapping indicators	Subject matter
3.6	Respect principles of sustainable development and regard for the environment
5.4	Existence of environmental responsibility policy and measures
5.1	Transparent process for allocating resources in declared non-profit objectives
5.2	Redistribution policy and programmes for main stakeholders
6.3	Adopt accounting control mechanisms and external financial audit (The definitions for the scores of 6.3 and 6.4 were mistakenly transposed in the questionnaire)
6.4	Adopt policies and processes for internal control

10.2.4 Clarification for specific indicators

Indicator	Topic	Note
2.2	Organisational structure and chart	The existence of an organisational chart is a technicality and quite a few IFs do not publish one. It seemed disproportionate for this single factor to make the difference between a score of 1 and 3.
	<i>Policy</i>	IFs were not automatically marked down for not having an organisational chart; rather the score awarded was a qualitative assessment of the information published.
4.4	Make public positions for election and appointment	It was hard to differentiate between scores 2 and 3.
	<i>Policy</i>	3 was awarded where there was a specific announcement about positions, not just an indirect reference in a Congress agenda etc.; 3 was also awarded when there was evidence that the IF advertised staff vacancies
4.5	Eligibility rules and due diligence	This is a complex area and it proved difficult to give an adequate assessment in a single score.
	<i>Policy</i>	The score awarded represented a qualitative assessment of the relevant rules.



4.7	Representation of stakeholders and gender equality	The question covered two distinct factors.
	<i>Policy</i>	Scores awarded gave credit as long as at least one of the two factors was covered.
5.6	Education programmes	External provision did not necessarily seem to be an advantage.
	<i>Policy</i>	Score was not automatically limited to 2 if there was no evidence of external provision; rather the score awarded was a qualitative assessment of the evidence of provision.
5.7	Gender and geographical representation for solidarity programmes	The question covered two distinct factors.
	<i>Recommendation</i>	Scores awarded gave credit as long as at least one of the two factors was covered.

10.2.5 Level of agreement between self-assessed and moderated scoring

In order to test the level of agreement between the scores provided by the IFs and the moderated scores produced by I Trust Sport across 50 indicators of good governance (the first eight – from Section 1 on Guiding Codes are excluded), the responses were grouped into three categories. The first category included the indicators which were not fulfilled or only partially fulfilled, scoring 0 or 1; the second category comprised the indicators which were fulfilled, scoring 2; and the third category incorporated those indicators which were well-fulfilled or totally fulfilled (3 or 4).

Two measures of agreement were analysed to compare the self-assessed and moderated scoring:

- (i) Level of agreement: the percentage of responses where the moderated category was identical to the category assigned by the IF;
- (ii) Cohen's Kappa, a common indicator of inter-coder reliability that is used in research. In brief, the value of Cohen's Kappa ranges from 0 to 1 and the higher it is, the stronger the reliability between the coders' particular scorings. The following table provides an interpretation of how strong an agreement is based on the Cohen's Kappa value:

Interpretation of level of agreement based on Cohen's Kappa value

Value of Kappa	Level of agreement	% of data that are reliable
0-.20	None	0 - 4%
.21 - .39	Minimal	4 - 15%
.40 - .59	Weak	15 - 35%
.60 - .79	Moderate	35 - 63%
.80 - .90	Strong	64 – 81%
Above 90	Almost perfect	82 – 100%

Cited from McHugh (2012); more details about Cohen's Kappa can be found at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/>

These measures of agreement were tested against two samples. The first sample included all 28 IFs. The second sample was composed of 26 IFs, with the two biggest outliers excluded. These are two IFs whose self-assessment score was very significantly different from the total



moderated score (variances of +81 and +66) and could therefore disproportionately distort the findings. The following results were calculated for the inter-coder reliability in the two samples:

Sample	Level of agreement	Cohen's Kappa
All 28 IFs	73% (1020 cases of identical coding and 378 different)	0.585 Between weak and moderate agreement
26 IFs, excluding two outliers	76.3% (991 cases of identical coding and 307 different)	0.635 Moderate agreement

The findings above suggest a moderate level of agreement between the adjusted and self-assessed scores. Considering the IF representatives who completed the questionnaires and the I Trust Sport consultants who conducted the moderation, there were differences in the profile of the people involved, probably in the time allocated to the task, and perhaps even in the incentives. There was also a lack of detailed guidance for IFs regarding the coding of the individual indicators. It can therefore be concluded that the inter-coder reliability has been at a satisfactory level.



10.3 Credits and acknowledgements

Our sincere thanks go to the 28 International Federations that completed the governance questionnaire between November 2016 and February 2017. Without their support this report would not have been possible.

This report was written in co-operation with the ASOIF Governance Taskforce.

Chair:

- Francesco Ricci Bitti (Chair, ASOIF President)

Members:

- Patrick Baumann (FIBA Secretary General, IOC Member and ASOIF Council Member)
- Jean-Loup Chappelet (Professor at IDHEAP Swiss Graduate School of Public Administration at the University of Lausanne)
- Benjamin Cohen (Director, European Regional Office and International Federation Relations, WADA)
- Pâquerette Girard Zappelli, (Chief Ethics and Compliance Office, IOC)
- Luis Fernando Lima (FIVB Secretary General)
- Alexander McLin (Executive Director, ASA – Swiss Arbitration Association)
- Denis Oswald (FISA Honorary President, IOC Member and CIES Director)
- Andrew Ryan (ASOIF Executive Director)
- Ingmar de Vos (FEI President)

Consultancy support:

- Rowland Jack and Slobodan Tomic from I Trust Sport reviewed the questionnaire responses, moderated the scores and produced analysis for this report.

Thanks to the ASOIF staff for their help with the project.

No part of this publication may be copied, republished, stored in a retrieval system or otherwise reproduced or transmitted, in any form or by any means whatsoever, without the prior written consent of the Association of Summer Olympic International Federations (ASOIF). This publication and its contents are the property of ASOIF.

© ASOIF Lausanne, March 2017

ASOIF - Association of Summer Olympic International Federations
Maison du Sport International, Av. de Rhodanie 54, 1007 Lausanne - Switzerland
Tel.: +41 21 601 48 88
Fax: +41 21 601 48 89
Email: info@asoif.com

www.asoif.com