8. STATISTICAL METHOD FOR PROCESSING SCORES

8.1. The CIVA Fair Play System - Purpose

8.1.1.1. Calculation of grades and scores for an aerobatic competition Programme using a mathematical process to give equal importance to all judges, while replacing anomalous grades with statistically fitted values.

8.2. Overview

The rating of a pilot performance for a given flight is an amount of points arising from two separate sources:

8.2.1.1. An evaluation of the quality of flown figures and of a flight's positioning with a grade given by judges observing the flight, on a scale ranging from 0 to 10 in increments of 0.5. These grades are multiplied by difficulty coefficients for each figure and added to derive a score for the programme for each pilot.

8.2.1.2. Penalties arising from height or time infringements and/or interruptions of the program sequence and other disciplinary actions.

8.2.1.3. The scores from 8.2.1.1 are subject to random and systematic errors due to the inevitable lack of exactness of judging. The purpose of the Fair Play system is to reduce the effect of those errors to a minimum. The penalties from 8.2.1.2 are not subject to the same errors and are simply subtracted from the scores results 8.2.1.1 after they have been calculated as described below.

8.3. Pre-Processing

8.3.1. Dealing with Hard Zeroes and Missed Figures

8.3.1.1. Prior to the scoring data being entered into the computer, the Chief Judge must ascertain the validity of Hard Zero grades. If a figure is determined to have been a Confirmed Hard Zero, this must be designated by the Chief Judge. The grades given to that figure by the grading judges must not be altered prior to being input into the computer.

8.3.1.2. For a figure determined not to be a Confirmed Hard Zero, any "HZ" grade given by a grading judge must remain unaltered prior to data entry into the scoring computer.

8.3.1.3. Figures that have been missed by a grading judge must be marked "A". These missing grades will be replaced automatically by the Fair Play system.

8.3.2. Identifying Figure Grades for Analysis

8.3.2.1. Prior to the start of the Programme, the data input to the scoring computer will include the number of pilots, the number of figures (including positioning and, for gliders, harmony) each pilot will fly, the K-factors of each figure and the number of judges for the programme.

8.3.2.2. Each figure will be given a full identifying number in the format ***kkkffpp*** where:

a) *kkk* is the K-factor, with leading zeroes if necessary, e.g. "037" if K-factor = 37

b) *ff* is the figure number, with leading zeroes if necessary, and

c) *pp* is the pilot number, with leading zeroes if necessary.

8.3.2.3. Note that the number pp allocated to a pilot must remain the same throughout a contest and should not be confused with the flight order number any pilot may be allocated for a particular programme.

8.3.3. Grouping Figure Grades for Analysis

8.3.3.1. Statistical manipulation must only be carried out on sets of data of reasonable size. Furthermore, such analysis is best conducted on sets of data that share similar source characteristics. To meet these requirements, the grading data from a programme must be combined into appropriate groups.

8.3.3.2. For the purpose of the Fair Play analysis, data will be arranged in groups in the following generalised format:

K-factor	Figure #	Pilot #	Judge 1	Judge 2	…	…	Judge j
Kkk_1	ff	pp	$Grade_{1,1}$				
Kkk_2	ff	pp					
Kkk_3	ff	pp					
…	…	…					
…	…	…					
kkk_n	ff	pp					$Grade_{n,j}$

8.3.3.3. In such a data set, the arrangement of rows will be by ascending value of the full figure Identification Number *kkkffpp*. In compulsory programmes, Known and Unknown, all pilots fly the same figures and the number of rows per data group will normally be the same as the number of pilots. This means that each data group in a compulsory sequence will correspond to a figure of that sequence flown by all pilots, in the form:

K-factor	Figure #	Pilot #	Judge 1	Judge 2	…	…	Judge j
kkk	Figure 1	Pilot 1	$Grade_{1,1}$				
kkk	Figure 1	Pilot 2					
kkk	Figure 1	Pilot 3					
…	…	…					
…	…	…					
kkk	Figure 1	Pilot p					$Grade_{p,j}$

8.3.3.4. Exceptionally, if the number of pilots is less than 11 (see 1.2.3.3), the data will be sorted in increasing value of the K-factor and divided into groups as follows:

Number of Pilots	*2*	*3*	*4*	*5*	6	7	8	9	10
Group Size	*12*	*12*	*12*	*15*	18	14	16	18	20

8.3.3.5. In Free Programmes, where pilots fly different figures and/or numbers of figures, additional information is required so that the figures included in each data group are reasonably similar in type and complexity. Therefore each figure in a Free Programme (including Positioning and Harmony grades) will additionally be allocated to a Super-Family. Super-Families are defined as follows:

Super-Family Numbers (FF)	Unlimited Power	Advanced Power	Gliders
Harmony			00
Positioning	01	01	01
Aresti family 2	02	02	02
Figures containing spins	included below	03	03
Figures without spins but with flicks	included below	04	04
Aresti family 5	05	05	05
Aresti family 6	06	06	06
Aresti families 1, 7 and 8	07	07	07

Note: If either Super-Family 05 or 06 contains less figures than the minimum of 11 data points, these two Super-Families will be combined.

8.3.3.6. Hence a Full Free Figure Identification Number will be of the form FFkkkffpp.

8.3.3.7. Free Programmes.

a) In the Positioning and Harmony Super-Families, the group size will equal the number of pilots, i.e. each will contain the complete Super-Family. If the number of pilots (Np) whose flights have been judged is < 11, however, (see 1.2.3.3) then these Super-Families will be combined into a group containing them both.

b) In other Super-Families, comprising aerobatic figures, the data groups will be formed from within each Super-Family, unless Np is less than 11. The target number of rows for each group (NrmGrp) will be the number of pilots whose flights have been judged, while the minimum group size (MinGrp) will remain 11 rows. When Np < 11, then NrmGrp will be as tabulated in 8.3.3.4, and a group may contain figures from more than one Super-Family. When a Super-Family contains more figures than the number of pilots, it may thus be split into two or more groups.

i) The boundary between adjacent groups within a single Super-Family will be made preferably at the change of K-factor nearest the target size within the range ‘target row to target plus minimum rows’, or if this is not successful nearest the target size but between the target row and the minimum group size. If no change of K-factor is available the group boundary will be set at the target row.

ii) For example, suppose that a Free Programme has 40 pilots and that Super-Family 07 contains 250 figures. This data will be divided into a number of groups, each of which will contain approximately 40 rows. The final group will contain at least 11 rows.

8.3.4. Confirmation of Hard Zero

8.3.4.1. The first stage of processing is to set to “HZ” all numerical grades given to a figure subsequently deemed to be a Confirmed Hard Zero by the Chief Judge. Any grade thus reduced to “HZ” must result in an increment to the particular judge’s record for determining the HZI component of the Judges Performance Index.

8.3.4.2. Once Confirmed Hard Zeroes have been implemented, each pilot’s score sheet should be printed and made available for inspection along with the judges grading sheets.

8.3.5. Treatment of Other "HZ" or of "A" Grades

8.3.5.1. If a figure is not deemed to be a Confirmed Hard Zero, any "HZ" or "A" grades given for that figure must be treated as missing data points. Such grades will therefore be excluded from the calculation of means or standard deviations until such time as they are replaced later in the process. For each "HZ" grade that is not confirmed, an increment will be made to the judge's HZI.

8.3.6. Treatment of Soft Zero Grades

8.3.6.1. Soft Zero grades are not subject to the same confirmation process as Hard Zeroes. They are generally treated as valid numerical grades in the same way as non-zero grades. However, Soft Zero grades should not influence the normalisation of non-zero grades that is described below.

8.4. Definitions

8.4.1. The Basic Data Values

8.4.1.1. Define the Raw Grades, for a given sequence, as:
S(ff, pp, j)

This is the Grade awarded by Judge *j* to Pilot *pp* flying Figure *ff.*

8.4.1.2. These Grades are then divided into semi-homogeneous Groups as defined above, and are now defined as:
R_g(fp, j)

This is the Grade awarded by Judge *j* to (Pilot *p* flying Figure *f*) in Group *g*, and is represented physically by a rectangular array of numbers where *fp* is the row index and *j* is the column index.

8.4.1.3. There should also be a count indicator of values *0* and *1* to indicate *0* for any SZ, *HZ* or *A* values. These are designated:
N_g(fp, j)

8.4.1.4. Counts

a) Pilot Count = No. Judges who score this pilot/figure combination

$$\mathbf{C_g(fp, *) = \Sigma_j \{N_g(fp, j)\}} \quad \mathbf{(1)}$$

b) Judge Count = No. Pilot/figures scored by this judge

$$\mathbf{C_g(*, j) = \Sigma_{fp} \{N_g(fp, j)\}} \quad \mathbf{(2)}$$

c) Overall Count = Total number of Scores

$$\mathbf{C_g(*, *) = \Sigma_{fp,j} \{N_g(fp, j)\}} \quad \mathbf{(3)}$$

8.4.1.5. Mean Values

a) Pilot Mean

$$\mathbf{mR_g(fp, *) = \Sigma_j \{R_g(fp, j)\}/C_g(fp, *)} \quad \mathbf{(4)}$$

b) Judge Mean

$$\mathbf{mR_g(*, j) = \Sigma_{fp} \{R_g(fp, j)\}/C_g(*, j)} \quad \mathbf{(5)}$$

c) Overall Mean

$$\mathbf{mR_g(*, *) = \Sigma_{fp,j} \{R_g(fp, j)\}/C_g(*, *)} \quad \mathbf{(6)}$$

8.4.1.6. Standard Deviations

a) Judge Standard Deviation

$$sdR_g(*, j) = sqrt([\Sigma_{fp} \{R_g(fp, j)\}^2 - C_g(*, j)*\{mR_g(*, j)\}^2]/[C_g(*, j) - 1]) \quad (7)$$

b) Average Judge Standard Deviation

$$sdR_g(*,) = \Sigma_j \{sdR_g(, j)\}/J \quad (8)$$

8.5. Group Processes

8.5.1. Normalisation of a Data Group

8.5.1.1. The first stage of the analysis is to Normalise the non-zero grades in the data group to give each judge's column of grades the same standard deviation. This will give equal importance to each judge's opinion. In the normalisation formula:

a) $Norm1_g(fp,j)$ is the Normalised grade to replace the Raw grade
$sdR_g(*,j)$ is the standard deviation for a judge's Raw grades in this group
$sdR_g(*, *)$ is the standard deviation for all the Raw grades in this group from all judges and,

$$Norm1_g(fp,j) = mR_g(*,*) + [R_g(fp,j) - mR_g(*,j)] * sdR_g(*,) / sdR_g(*,j) \quad (9)$$

8.5.1.2. If the result of formula (7) or (8) is zero, then formula (9) cannot be applied and the grades for this judge, or this group, should be unchanged. If the result of formula (9) is less than zero, then it should be set at zero.

8.5.1.3. Soft Zero (0.0) grades are excluded from this normalisation process because, for each judge, these form part of a second mode of distribution of raw grades. After the non-zero grades are normalised, the Soft Zero grades are set at 0.0 so that they are included in the process of determining Fitted Values and figure anomalies Hence:

$$\text{If } Rg(fp,j) = SZ, \text{ Then } Norm1g(fp,j) = 0.0 \quad (9a)$$

8.5.2. Derivation of Fitted Values

8.5.2.1. Within the data group, a Fitted Value for a figure grade for a pilot is the grade that you would expect a particular judge to give a particular pilot/figure combination, based on an analysis of all the judges' grades for all the pilot/figure combinations in the group, including numerical zeroes (SZ) but excluding factual zeroes (HZ). In the Fitted Value formula:

a) $FV1_g(fp,j)$ is the Fitted Value derived from $Norm1_g(fp,j)$
$mNorm1_g(*,j)$ is the mean of the Normalised numerical grades in the group for that judge
$mNorm1_g(fp,*)$ is the mean of the Normalised numerical grades in the group for that pilot/figure
$mNorm1_g(*, *)$ is the mean of all the Normalised numerical grades for that group for all judges and,

$$FV1_g(fp,j) = mNorm1_g(*,j) + mNorm1_g(fp,*) - mNorm1_g(*, *) \quad (10)$$

8.5.3. Assessment of Anomalous Grades

The normalised grades in each group must be tested for anomalies caused by judging error or partiality.

8.5.3.1. The Uncertainty of Any Individual Data Point

a) A data point (grade) will be considered anomalous if its uncertainty exceeds a given threshold value. This uncertainty is derived by a two-way analysis of variance and starts with the calculation of the Residual for each data point. In the Residual formula:

b) $Res1_g(fp,j)$ is the Residual value for each data point in the group after the first normalisation, and,

$$\mathbf{Res1_g(fp,j) = Norm1_g(fp,j) - FV1_g(fp,j)} \quad \mathbf{(11)}$$

c) $RSS1_g$ is the Residual Sum of Squares for the data group after normalisation and,

$$\boldsymbol{RSS1_g = \sum_{fp,j}\{Res1_g(fp,j)\}^2} \quad \mathbf{(12)}$$

8.5.3.2. The Degrees of Freedom of the data group is determined by:

a) D_g is the value of the Degrees of Freedom of the data group
FP_g is the number of pilot/figure rows in the group
J_g is the number of judges in the programme (columns in the data group)
Nm_g is the number of missing values (HZ or A) in the group, and

$$\mathbf{D_g = \{[FP_g - 1] * [J_g - 1]\} - Nm_g} \quad \mathbf{(13)}$$

8.5.3.3. The Residual Standard Deviation of the data group, $RSD1_g$, is determined by:

$$\mathbf{RSD1_g = sqrt\{RSS1_g / D_g\}} \quad \mathbf{(14)}$$

8.5.3.4. Finally, the uncertainty of each individual data point, $U1_g(fp,j)$ is calculated:

$$\mathbf{U1_g(fp,j) = ABS[Res1_g(fp,j)] / RSD1_g} \quad \mathbf{(15)}$$

8.5.4. Treatment of Anomalous Grades

8.5.4.1. If the uncertainty of an individual grade, $U1_g(fp,j)$, exceeds 2.24 it has an uncertainty of approximately 97.5%. This degree of anomaly, or more, is to be expected in the case of a small number of soft zeroes for a figure which generally attracts a majority of high grades. Similarly, such an anomaly might occur if a single judge missed a large pilot error that led all other judges to award a very low grade. Anomalies such as this should be treated as though they were missing values. This treatment will give the benefit of the doubt to the pilot in situations where it is possible that a very significant judging error has been made.

8.5.4.2. The raw grade for any data point showing such an anomaly should be set to “Missing” in the original Raw Data Rg(fp, j) – call it R2g(fp, j). The judge concerned should have an increment made to his LSI or HSI component of the Judges Performance Analysis, as appropriate, for each grade replaced.

8.5.4.3. When making judgements based on the perception of the quality of flick rolls or spins, the panel of judges might produce a series of grades in which the distribution is bi-modal rather than Gaussian. For example, a set of grades might possibly include a number of soft zeroes and a number of high grades. In extremely rare cases, this difference of opinion may be so great that the majority of raw grades might be considered anomalous by this analysis. In this situation it is not fair to assume that the remaining grades are truly representative of the pilot’s performance of the figure concerned.

8.5.4.4. Therefore, if the number of missing values that would be carried forward to the second normalisation exceeds 60% of the number of judges, all grades for this figure by this pilot should be replaced by the FV1 value derived at Formula 10.

8.5.5. Second Normalisation of the Group

8.5.5.1. If anomalies have been removed from the raw grades, the data set will have more missing values. It would therefore be necessary to normalise the data group for a second time. Again, Soft Zero (0.0) grades must be excluded from the Normalisation and these grades

must remain 0.0. Using only the remaining non-zero grades, new values must be determined for $mNorm_g(*,j)$, $mN_g(fp, *)$, $mN_g(*, *)$ and thus $FV_g(fp,j)$.

a) Hence,

$$\mathbf{Norm2_g(fp,j) = mR2_g(*.*) + [R2_g(fp,j) – mR2_g(*,j)] * sdR2_g(*,) / sdR2_g(*,j)} \quad \mathbf{(16)}$$

b) and,

$$\mathbf{FV2_g(fp,j) = mNorm2_g(*,j) + mNorm2_g(fp,*) – mNorm2_g(*, *)} \quad \mathbf{(17)}$$

8.5.5.2. These new fitted values will have been determined free from the influence of any anomalous grades and are thus robust and give the benefit of any doubt to the pilot in the case of minority soft zeroes for an otherwise highly-graded figure.

8.5.6. Replacement of Missing Grades

8.5.6.1. These $FV2_g(fp,j)$ values are then used to replace the HZ, A and 'Missing' anomalous grades carried forward from the preceding analysis.

8.5.6.2. The judge concerned should have an increment made to his LSI or HSI component of the Judges Performance Analysis, as appropriate, for each anomalous grade replaced, as well as to the HZI component for any HZ replaced.

8.5.6.3. After these replacements, the second normalised grades will be the final processed grades for each data group.

8.5.7. Assembly of Processed Grades by Pilot

8.5.7.1. After processing in the separate data groups, the final processed grades must be combined into a single matrix and this table sorted by ascending value of the Pilot identification number and then the figure number. These grades are then multiplied by the respective K-factor for each figure and totalled to give:

a) SR(p,f,j) an overall score for each pilot on each figure from each judge

8.5.7.2. These can then give

a) SR(p,f,*) an overall score for each pilot for each figure over all judges, where:

$$\mathbf{SR(p,f,*) = \Sigma_j\ SR(p,f,j)} \quad \mathbf{(18)}$$

b) SR(p,*,j) an overall score for each pilot for each judge over all figures, where:

$$\mathbf{SR(p,*,j) = \Sigma_f\ SR(p,f,j)} \quad \mathbf{(19)}$$

c) SR(p,*,*) an overall score for each pilot, where:

$$\mathbf{SR(p,*,*) = \Sigma_{f,j}\ SR(p,f,j)} \quad \mathbf{(20)}$$

8.5.7.3. These data should be printed and passed to each pilot at the earliest possible stage, so that the changes made during the processing stage can be understood.

8.6. Sequence Processes

8.6.1. Normalisation of Sequence Scores

8.6.1.1. It is now necessary to repeat the normalisation process at the sequence stage, once again to ensure that the opinion of each judge is given the same importance.

8.6.1.2. The sequence score data, SR(p,j), can be set out in a matrix form as shown here.

Pilot #	Judge 1	Judge 2	Judge 3	…	…	…	Judge j
Pilot 1	SR(1,1)	SR(1,2)	…	…	…	…	…
Pilot 2	SR(2,1)	…	…	…	…	…	…
Pilot 3	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…
Pilot p	…	…	…	…	…	…	SR(p,j)

8.6.1.3. From this table:

a) mSR(*,j) is the mean of all the scores given by Judge j.
sdSR(*,j) is the standard deviation of all the scores given by Judge j.
sdSR(*,*) is the average standard deviation of all the scores given to all the pilots by all the judges, and

$$\mathbf{NormS(p,j) = mSR(*,j) + [SR(p,j) - mSR(*,j)] * sdSR(*,*) / sdSR(*,j)} \quad \mathbf{(21)}$$

8.6.2. Derivation of Sequence Fitted Values

8.6.2.1. Next sequence fitted values are derived from the normalised scores to enable calculation of standardised residuals at the sequence level. In this derivation:

a) mNormS(*,j) is the mean of all the normalised scores given by Judge j.
mNormS(p,*) is the mean of all the normalised scores given to Pilot p.
mNormS(*,*) is the mean of all normalised scores given by all judges to all pilots, and

$$\boldsymbol{FVS(p,j) = mNormS(*,j) + mNormS(p,*) - mNormS(*,*)} \quad \mathbf{(22)}$$

8.6.3. Assessment of Sequence Anomalies

8.6.3.1. Despite the replacement of anomalous figures at the earlier stage of the process, it might be possible for slight, consistent favouritism or subconscious bias to influence unduly a Judge's overall score for a pilot. Such a score might be high or low and should be replaced if its degree of uncertainty reaches approximately 90%.

8.6.3.2. Therefore the analysis must next derive the residuals for the sequence scores:

$$\boldsymbol{ResS(p,j) = NormS(p,j) - FVS(p,j)}\text{, and} \quad \mathbf{(23)}$$

$$\boldsymbol{RSS_s = \sum_{p,j}\{ResS(p,j)\}^2} \quad \mathbf{(24)}$$

8.6.3.3. The number of degrees of freedom for the sequence data set is calculated where:

a) D_s is the value of the Degrees of Freedom of the sequence data
P_s is the number of Pilots in the sequence
J_s is the number of judges in the sequence
Nm_s is the number of missing values (confirmed HZ for all figures by a pilot), and

$$\mathbf{D_s = \{[P_s - 1] * [J_s - 1]\} - Nm_s} \quad \mathbf{(25)}$$

8.6.3.4. The Residual Standard Deviation for the sequence is given by:

$$\mathbf{RSD_s = sqrt\{RSS_s / D_s\}} \quad \mathbf{(26)}$$

8.6.3.5. The uncertainty of each sequence score is given by:

$$US_s(p,j) = ABS[ResS(p,j)] / RSD_s \quad (27)$$

8.6.3.6. If this uncertainty figure exceeds 1.65 (90%) it must be replaced by the fitted value FVS(p.j).

8.6.4. Interim Final Sequence Score

8.6.4.1. The processed sequence score for each pilot will be the sum of the normalised sequence scores over judges, after replacement of anomalous values of NormS(p,j) by fitted values FVS(p,j).

$$PS(p) = \sum_j \{NormS(p,j) \text{ or } FVS(p,j)\} / Nj \quad (28)$$

8.6.4.2. Penalties awarded for whatever reason are subtracted from this processed score to give each pilot's final overall score for the sequence.

$$FS(p) = PS(p) - Pen(p) \quad (29)$$

8.6.5. Second FPS Iteration and Final Sequence Score

8.6.5.1. When flights are of a very low standard, it is unlikely that the judges will show the same consistency of grading as when flights are of a high standard. Therefore, such low standard flights can have undue influence over the way in which the FPS system treats other scores.

8.6.5.2. To prevent such undue influence, the following procedure will be followed if the total number of competing pilots exceeds 30:

a) Determine the values of PS(p) as a percentage of the maximum possible score for the sequence.

b) If this value is less than 60% for a known sequence, or less than 50% for an unknown sequence, temporarily remove these flights raw data from the whole data set and re-apply the FPS process in its entirety. This will generate more reliable results for the retained pilots.

c) Publish the final ranked order, based on FS(p) from the first FPS iteration for the excluded, low-scoring pilots, and based on FS(p) from the second FPS iteration for the retained, higher scoring pilots.

8.7. Process Summary

8.7.1. The process carries out the following analytical steps:

8.7.1.1. Sets confirmed Hard Zeros to HZ for all judges

8.7.1.2. Treats unconfirmed HZ and A grades as "Missing" at this stage.

8.7.1.3. Arranges figure grades into data groups for further analysis.

8.7.1.4. Within each data group:

a) Normalises the grades to give equal importance to each judge.

b) Derives fitted values for each judge for each figure.

c) Determines if any normalised grades are more than 97.5% uncertain and disregards them by setting them to "Missing".

d) Derives revised normalised grades and fitted values taking account of the new missing data.

e) Replaces all the missing grades with revised fitted values.

8.7.1.5. At the sequence level:

a) Normalises the scores to give equal importance to each judge.

b) Derives fitted values for each judge for each pilot.

c) Determines if any scores are more than 90.0% uncertain and replaces them with fitted values.

8.7.1.6. In the Second Iteration:

a) Repeats the FPS process excluding certain low-scoring flights.

b) Recombines all results into a final ranking order.

8.8. Judging Performance Indices

The JPI system generates judging analysis data from the raw and FPS-processed scores. Six different aspects of judging performance are studied and each gives rise to its own index which is independent from the number of sequences and figures flown in a particular programme. The six individual indices are described below. In each case, the lower the derived value of the index, the better is the performance of that individual judge.

8.8.1. Ranking Index (RI)

8.8.1.1. The Ranking Index measures how closely an individual judge's pilot ranking for a programme conforms to the overall ranking based on all judges' assessments.

8.8.1.2. For each judge, determine for each pilot the difference between the overall ranking R and the judge's ranking Rj. Sum all these differences and then divide by the square of the number of pilots to get an index that is independent of field size. If there are N pilots in the programme, then:

$RI = \frac{\sum_{1}^{N} \sqrt{(R - Rj)^2}}{N^2} x2$. Typical values are between 0.05 and 0.25, maximum 0.5.

8.8.2. Low Scoring Index (LSI)

8.8.2.1. The Low Scoring Index measures how many times a judge grades a figure significantly lower than the consensus view of the judges.

8.8.2.2. For each figure, examine the normalised scores. If a judge's score for the figure has been determined 'Low' at the approved confidence level, then add one to that judge's aggregate of errors (EL) under this heading. When all figures for all pilots have been graded, divide the judge's sum of errors by the total number of figures observed.

8.8.2.3. If the number of competing pilots is P and the number of figures in the sequence is F, then:

$LSI = \frac{\sum E_L}{PxF}$. Typical values will be between 0.04 and 0.2.

8.8.3. High Scoring Index (HSI)

8.8.3.1. The High Scoring Index measures how many times an individual judge grades a figure significantly higher than the consensus view of the judges.

8.8.3.2. For each figure, examine the normalised scores. If a judge's score for the figure has been determined 'High' at the approved confidence level, then add one to that judge's aggregate of errors (EL) under this heading. When all figures for all pilots have been graded, divide the judge's sum of errors by the total number of figures observed.

8.8.3.3. If the number of competing pilots is P and the number of figures in the sequence is F, then: $HSI = \frac{\sum E_H}{PxF}$. Typical values will be between 0.02 and 0.1.

8.8.4. Discrimination Index (DI)

8.8.4.1. The Discrimination Index measures the range of raw scores being used by an individual judge to differentiate between well-flown and poorly-flown figures

8.8.4.2. Count the number of times during the whole programme that an individual judge uses each of the non-zero raw scores of 0.5 to 10.0. Calculate the population variance (VARp) for this data set. Divide this variance by two and then subtract the result from one to get the Discrimination Index.

8.8.4.3. Thus: $DI = 1 - \frac{VARp}{2}$. Typical values will be from 0 to 1. Negative values are possible, but these should be treated as zero (If DI < 0, then DI = 0).

8.8.5. Hard Zero Index (HZI)

8.8.5.1. The occurrence of Hard Zeroes is determined by majority voting or by video conference. The scoring system determines the application of the Index from the “CHZ” box on the score sheets.

8.8.5.2. In the event that an individual judge fails to identify a confirmed hard zero, then add one to that judge’s aggregate of errors (EZ) under this heading. Similarly, if a judge gives a grade of HZ when no such error occurred, add one to the aggregate of errors (EZ) under this heading.

8.8.5.3. In the unusual case that the aggregate of Hard and Soft Zeroes represents a majority of the judges’ opinions, and the figure is subsequently confirmed as a zero, then no HZI error penalty will be awarded to any judge who initially graded the figure HZ or 0.0.

8.8.5.4. If the number of competing pilots is P and the number of figures in the sequence is F, then: $HZI = \frac{\sum E_Z}{PxF}$. Typical values will be between 0.0 and 0.05.

8.8.6. Sequence Anomaly Index

8.8.6.1. The Sequence Anomaly Index measures how many times a judge grades a whole sequence significantly higher or lower than the consensus view of the judges.

8.8.6.2. For each pilot, examine the normalised sequence scores. If a judge’s score for the sequence has been determined ‘High’ or ‘Low’ at the approved confidence level, then add one to that judge’s aggregate of errors (ES) under this heading. When all figures for all pilots have been graded, divide the judge’s sum of errors by the total number of sequences observed

8.8.6.3. If the number of competing pilots is P, then: $SAI = \frac{\sum E_S}{P}$. Typical values will be between 0.0 and 0.1.

8.8.7. Overall Judging Performance Index (JPI)

8.8.7.1. It is possible to combine the results of the different index calculations into one overall Judging Performance Index that is independent of the number of judges in the panel.

However, the Discrimination Index should not be included in this overall figure as it relates primarily to style, not to accuracy.

8.8.7.2. For each of the remaining five separate indices, each judge is given a ranking from 1 (best) to N (the Number of judges). These rankings are then added for each judge. The sum of these additions is also calculated and divided by the number of judges, to give a mean ranking score. Each judge's personal ranking total is then divided by the average to get an overall JPI that will average unity among all the judges.

8.8.7.3. In any particular corps of judges, the better individuals will have a JPI less than 1, while those performing less well will have a JPI exceeding 1. The further these individual scores are from unity, the greater is the difference in judging skill between the best and the worst, for any particular programme.