

Report of the CIVA FPS Working Group 2015

The CIVA FairPlay System Working Group had maintained a watching brief during 2015 to monitor operational aspects of the algorithm during the year's championships, the work of the previous year having concluded that no changes of any significance were required.

The following aspects of FPS have however received closer attention from the group members this year than has been the case in the past:

Smoothness of reaction to data confidence changes

During 2015 particular attention has been paid to the smoothness of output of the system at the margins, that is to say with regard to small but discernible step changes that might be observed as a consequence of the statistical grade and score accept / reject process.

Whereas human confidence in an event almost always moves linearly from complete confidence to none (or vice versa) as the underlying data changes, classical statistical confidence testing theory sees this process as an abrupt change from full confidence, i.e. acceptance that a grade or score fits the pattern of other data, to no confidence, i.e. rejection because a grade or score is deemed not acceptable by comparison with other data; this step occurs at whatever value the confidence SD is set to. When these step changes occur in FPS they are small and generally lost in the far wider picture of the whole sequence, nevertheless the view has been expressed by some that moving the FPS grade and score acceptance / rejection process to one that more closely follows human 'linear' thinking could be a desirable refinement.

An FPS group discussion document has thus been circulated recently that calls for comment in this and other areas, and we expect it either to encourage and lead to further investigation or draw this line of thought to a close.

Review of the construction of the judges Ranking Index

A judge who awards a wrong HZ or misses one that should have been applied is likely to significantly damage his RI because this failure represents a moment when the judge should have – but clearly was not – paying sufficient attention. If the pilot is toward the bottom of the rankings however the calculated RI element for a missed HZ can be less than for a higher ranking pilot because there is less room for the pilots rank to be depressed, even though the error by the judge is effectively the same.

It is the view of many Chief Judges that missing HZ's or giving them in error is the most serious mistake that a judge can make, and if this happens too often then the effect on the judges RI should be numerically larger than is currently the case.

The FPS group discussion document mentioned above thus examines how we could emphasize the awarding of too many 'wrong' HZ's by making their impact on the judges RI more significant if the number given rises above a fixed number.

Possible outcome of the above discussion process

Should either of the above discussion areas lead to recommendations for change then the FPS working group will make a proposal for consideration by the 2016 plenary and provide detailed evidence of the processes under consideration.

A handwritten signature in blue ink, appearing to read 'Nick Buckenham', is written in a cursive style.

Nick Buckenham
CIVA FPS WG Chairman 2015

For reference:

Judge Ranking Indexes can be derived from any set of results – whether using FPS or unchanged raw grades – and are a direct measure of how similarly (or not) each judge has ranked the pilots when compared to the final results before penalties are applied. An RI of zero is perfect (the judge has ranked all of the pilots identically to the panel) whereas larger RI numbers are evidence of more marked differences between the judge's opinions and those of the panel. In general when FPS is used an RI below 10-15 for a single sequence is acceptable, but if this rises to over 20 it indicates increasingly significant differences of opinion between the judge and the panel that should be further investigated.

At CIVA events the "RI" for each judge is built by comparing the order of pilot ranking from each individual judge against the final results order after FPS processing is completed but before penalties are applied. When there is a difference between a judges' ranking of a pilot and the post-FPS / pre-penalties results this rank difference is factored with the difference between the judges' score and the final score to provide a Ranking Index element for that pilot, and this is done in a way that seeks to avoid the influence of the number of pilots in the data pool. Both ranking and score differences per pilot thus contribute to the impact on the judges RI, these contributions being totalled for each judge to provide the RI numbers that we see at each event.