TBL: Villian or Saint. Part Two

The 2005 F3A World Championship is using the TBL scoring mechanism for discarding judge scores. The scoring software was developed by Alexandre Pignot, and has received approval from the CIAM Bureau for application at this world championship. It is not the first time that the TBL process has been applied at a world or continental championship. A similar scoring process is also being used in full size aerobatics for well over two decades. The following report briefly outlines the steps involved in applying the TBL scoring mechanism. As there is a significant mathematical content involved in obtaining a TBL ranking, it is best to illustrate how the system operates by way of an example.

When TBL is being applied, the judges' raw scores for every competitor (from the competitor in first place to the competitor in last place) counts towards the round score awarded to the competitors. This process involves a number of steps:

Step 1: Standardise the Scores.

This phase modifies the score each judge awarded to each competitor. Judges scores may be transformed either up or down with the transformed scores maintaining the same competitor ranking as before. This modifying of judges raw scores attempts to remove any natural bias within a score. The standardisation process is applied as follows:

An average score and standard deviation is obtained for the whole panel. Consider the following example of six competitors and five judges. This gives a panel of 30 individual scores as illustrated in Table 1.

	Judge 1	Judge 2	Judge 3	Judge 4	Judge 5
Comp. 1	476	530	450	490	483
Comp. 2	520	520	497	518	545
Comp. 3	415	480	428	444	436
Comp. 4	378	490	415	385	396
Comp. 5	260	281	240	253	278
Comp. 6	65	120	72	80	78

Table 1 Raw scores awarded by the panel of judges.

The average score for the panel is **367.43**, with a standard deviation of **155.51**

The average score and standard deviation for each judge is then calculated. Results for this calculation are shown in Table 2.

	Judge1	Judge2	Judge3	Judge4	Judge5
Average	352.3	403.5	350.3	361.6	369.3
Standard Deviation	166.7	166.3	162.0	166.9	168.6

Table 2 Judges' Average Score and Standard Deviation

Taking these averages and standard deviations along with the average and standard deviation for the whole panel, each raw score shown in Table 1 is now transformed. After this transformation, the raw scores of each judge will have the same mean and standard deviation as that of the whole panel. The transformed results are shown in Table 3.

	Judge 1	Judge 2	Judge 3	Judge 4	Judge 5
Comp. 1	482.7	485.7	463.0	486.9	472.2
Comp. 2	523.7	476.3	508.1	513.0	529.4
Comp. 3	425.8	438.9	441.9	444.1	428.9
Comp. 4	391.3	448.3	429.4	389.1	392.0
Comp. 5	281.3	252.8	261.5	268.0	283.1
Comp. 6	99.4	102.3	100.3	105.0	98.7

Table 3 Judges' transformed scores

A look at the adjusted scores for competitor 6 gives a good indication of how the TBL system has corrected the imbalance in the original raw scores.

Step 2: Discard unwanted scores

Once the table of transformed scores has been obtained, the process of dropping judges' scores can begin. If any of the judges' scores for a competitor falls outside a set range, those judges' scores are discarded, and the process is repeated until no further scores are discarded. The competitor is then awarded the average of the remaining scores. Taking the results from Table 3, high and low limits are established and the competitor's final score is calculated as shown in Table 4.

	Judge 1	Judge 2	Judge 3	Judge 4	Judge 5	Low Limit	Upper Limit	Score
Comp. 1	482.7	485.7	463.0	486.9	472.2	461.2	494.9	478.1
Comp. 2	523.7	476.3	508.1	513.0	529.4	476.1	544.1	490.1
Comp. 3	425.8	438.9	441.9	444.1	428.9	422.5	449.2	435.9
Comp. 4	391.3	448.3	429.4	389.1	392.0	362.8	457.7	410.2
Comp. 5	281.3	252.8	261.5	268.0	283.1	248.0	290.6	269.3
Comp. 6	99.4	102.3	100.3	105.0	98.7	92	105.2	101.1

Table 4: Final table showing high and low limits along with the competitors' round score

In this example, none of the judges' scores will be dropped.

Step 3: Normalise Scores

When a final score is reached for each competitor, normalisation occurs as with previous scoring systems. The highest scoring competitor of the round is awarded 1000 points, and all other scores are adjusted accordingly.

NOTES:

1) If a competitor happens to fly only a few manoeuvres he may end up with a negative score, but it is not very likely. This is caused by the judge's scores being adjusted downwards and the adjustment factor being larger than the raw score for the competitor. Although it doesn't look natural on a score sheet, it is a valid TBL ranking.
2) Using the TBL system, we are 90% certain that any bias apparent in the raw scores is removed from the transformed scores,
3) Unlike previous scoring systems whereby a high/low judge or manoeuver was thrown out, TBL will only discard scores that are totally outside a set range. In front of a panel of 5 judges, there may be zero, one or two scores dropped.
4) TBL will retain on average 90% of all scores and perhaps discard only 10% of scores. The high-low throwaway system discards 40% of all available data, and retains only 60% of scores.
5) Dropped scores may be high, low or a combination of high and low scores.
6) Unless a competitor has an in-depth knowledge of how TBL operates, along with all the other competitors' raw scores from each judge, it would be impossible to determine how they would have fared if they had scored higher points in particular manoeuvres.

(originally written by David Power).