

# FPA Judging System Changes – Explanations

*February 4<sup>th</sup>, 2014*

## 1. Artistic Impression

**Flow:** Will remain the same, except becomes a subcategory that stands alone (see Summing Up section of this document); not be split into individual flow and team flow, because this split would make judging more (not less) complicated. Moreover, it would be perceived to deemphasize the importance of Flow, which should have elevated importance, according to players' feedback.

**Show:** Music Choreography, Individual Flow and Overall Impression were proposed to become part of the new category called "Show". This idea was disapproved by many players, because they don't like the connotation of the word "Show" and concerns about reducing styles of play down toward only crowd-pleasing moves. Moreover, it is seen as opening the possibility of too much subjectivity with judges giving more points to the teams they like. So no new category emphasizing "Show" will be implemented.

**Overall impression:** Will be abolished because it is totally subjective and this category influences all areas of judging anyway.

**Variety:** Several modifications were discussed:

1. Some players argued that Variety is mainly a technical aspect and should be part of Difficulty. The committee strongly considered this idea but finally refused it, because Diff judging is highly demanding already and focussing on Variety while also scoring Difficulty would clearly overstrain these judges.
2. Another idea was to give Variety scoring to Execution judges, but this seems strange because Execution and Variety are not related at all.
3. Having Variety as a fourth category with separate judges was also discussed, but since judges are a scarce resource at tournaments, this is not pragmatic.

Conclusion: Variety will stay a part of AI. However, to get more objective scores, judges will be handed a **Variety checklist**, which they should look at after each team's routine to see how many different areas of Freestyle play were attempted:

- The checklist will work as structure guideline and a support tool.
- The checklist groups the elements of Freestyle into 5 subcategories (throws, catches, disc handling, styles of play, spins/ambidexterity) each of which will get a subscore of 0-2, summing up to a maximum total variety score of 10. Guidelines how to allocate the 0-2 scores are part of the variety checklist.
- Within the subcategories, there will be no static rules (e.g., if a judge checks 7 catches the Catches subscore should be a 1.5).
  - o Static rules are not pragmatic here, firstly because applying them correctly will consume too much time during tournaments.
  - o Secondly, the elements of Freestyle are infinite and such a list (e.g. of catches) can never be complete. So it will contain the main/standard range of elements per subcategory only.
- Giving scores per Variety subcategory means – in contradiction to the old judging system – that poor variety in e.g. 'styles of play' cannot be fully made up for by a high variety in other Freestyle elements like 'throws' or 'catches'. This step was taken to incentivise players to show a high variety in all realms of Freestyle Frisbee and counteract the often criticised homogenisation of new school Freestyle play.

- The general idea of the variety checklist is that judges look at the list, quickly evaluate the various types of tricks demonstrated by the team, take the additional techniques shown into account and calculate their 5 subscores. This should be a rather fast, intuitive process in order to avoid further slowdowns of tournament progress.

**Form, Teamwork, Music Choreography:** Will remain the same.

**Summing up:** Five of the existing six AI categories will remain:

1. Variety,
2. Teamwork,
3. Music Choreography,
4. Form,
5. Flow.

Each of these categories will be judged from 1-10 points.

As Flow and Form used to be 1-5 points only, they are scored with more weight. Because the new proposed AI category “Show” was dismissed, this can be seen as a compromise and concession for the players who want visually compelling styles of play to be more incentivised.

One could argue that dismissing just one category (General Impression) is not a real simplification of AI judging, but throughout extensive discussions, it was again clear that Freestyle Disc is a sport with many facets, with many complexities of which – if overly simplified - will no longer be judged accurately.

## 2. Difficulty

The **new Diff tape (hybrid approach)** was welcomed and will be implemented. 3, 4 and 5-minute versions are available for download.

The **multiplier** was also generally approved of by FPA members given the hard facts we presented. Some people asked why we are using a static multiplier of 1.5 instead of an “exact” one that is balancing the variance of all categories. The exact multiplier changes the scores of all categories by defining the best team’s score of a category as 10 and multiplying all other teams’ scores by the same factor (exact formula available on request). This is mathematically more complex but feasible, since most tournaments have computer laptops at the events.

The committee tested and thoroughly discussed both multiplier options (static and exact) and decided in favor of the **static** one for two reasons:

1. First, the exact multiplier brings in a “black box” (competitors not being able to track calculations) to judging. At some point you leave the calculation of scores to a computer and many people will have trouble understanding how their score was computed.
2. Second, an exact multiplier doesn’t necessarily address the systematic lack of variance of Difficulty scores because it doesn’t always do the desired thing (i.e., it increases the variance of all categories with low variance in a given pool of teams even if there has been little difference between the teams in a category in reality). For example, one might make a small AI scoring gap between two teams bigger using an exact multiplier, although the small scoring gap perfectly reflects their AI performance in the run.

Note: The multiplier is implemented as a **transitional measure** only until judges are better educated and use the full range of Diff scores. As a first measure to get there, we will **add a legend to each Diff scoring sheet**, (i.e., a guideline for judges how to allocate their Diff

scores: 0: no tricks shown; 1-2: very easy tricks; 3-4: easy tricks; 5-6: medium tricks; 7-8: difficult tricks; 9-10: very difficult tricks). The idea is that people would rather say that something is very difficult than scoring it 9 or 10. But if they see that ‘very difficult’ should be scored 9 or 10, they are more likely to give those scores. We think that we can’t do much wrong by implementing this.

### 3. Execution

Basing Execution deductions only on the degree of breaks in flow that judges perceive during the routines was considered too subjective and a ‘nightmare to implement’. However, the idea of giving **greater** emphasis to the possibilities of **reducing execution penalties if an error does not influence the flow of play significantly** was welcomed. Based on players’ feedback we developed the following guideline:

.5 - for severe drops (throwaways; endangering crowd)
.3 - for real misses of the disc not touching the player's hand and interrupting the flow significantly (applies not only for catch attempts, but also missed pulls, brushes, etc.)
.2 - for drops that touch the player's hand or drops that do not touch the player’s hand but the disc is brought back into play without interrupting the flow significantly, e.g. the player immediately picks the disc up and brings it back into play
.2 - for unintended 'the' catches (seems subjective, but realistically all ‘the’ catches from players who are not total beginners can be seen as unintentional)
.1 - for all other execution mistakes like wobbles/bobbles, multiple 'the' brushes in a row, unclean roles, etc.

The possibility of handling Execution mistakes like this has been within the “wobble room” of the current judging system already, but we want to make clearer now that good Execution is not just about the number of drops but also about the overall flow of the presentation. Not every drop is a .3 and a “save” (catch) is not always just a .1 deduction.

### 4. The Bonuses (Uniqueness of Play, Speed Flow, Consecutivity)

The **Bonuses for Uniqueness/Creativity of Play, Speed Flow and Consecutivity will not be implemented in the proposed form**. The weight of bonuses (1.5 points in total per category) is seen as too high especially given the fact that judges can decide on them arbitrarily without any accountability for why they give the bonus points to any particular team. This is likely to lead to inflated scoring (pushing the team the judge likes) and strong disputes. Members commented that the proposed bonus point category of “Uniqueness/creativity” was deemed too subjective and a big part of current AI judging anyway.

Speed Flow and Consecutivity, however, are seen as important Freestyle elements, that many feel are not valued enough by the current judging system (even though they are part of it already). Other ways of making these two areas of skill more prominent/important should be further discussed.

- For Speed Flow we propose that this should be done, first, through judging education with an emphasis on the higher difficulty level of Speed Flow. **Speed Flow is more difficult than it appears** because it contains many catches; and more catch attempts always carry a higher risk of dropping the disc. Second, judges should be reminded to **make use of the possibility to give .2 deductions for drops occurring during a**

**speed flow, given the overall flow within a series of throws and catches** (the context of Speed Flow).

- Consecutivity of play has been part of Diff judging already. However, to encourage Consecutivity to be better acknowledged, Diff judges will be asked to note down a '+', a '-', or a '+-' together with each time block. This will **indicate good, poor or mediocre Consecutivity** during this time block which will influence their Diff score. When giving a '+' judges should increase their time block score by 1; a '-' should lead to a 1 point reduction; and a '+-' to no change. The idea is to continuously compel judges to consider Consecutivity with a simple system of taking it into account. This is supposed to be a transitional measure until the concept of Consecutivity is internalized and properly valued by all judges (again through education and experience).

To adequately judge Speed Flow and Consecutivity it is important to **give more prominence to both concepts during judging clinics and in the judging manual**. The committee has therefore written down explanations and examples that should become part of an Appendix of the FPA judging system. Moreover, we recommend to watch the sections of the Secrets of Pro Disc Freestyle video produced by Dave Lewis and Z Weyand (Consecutivity is called Connectivity here).

## 5. Crossing out high and low scores for AI and Diff

After the last round of players' input on Shrednow, the committee discussed again the ideas of having 5 judges for AI and 5 judges for Difficulty, and eliminating the high and low scores per category (i.e., for each team the best and the worst AI and Diff judging score would not be calculated for the average score). This would minimize the biasing effect of outliers, reduce subjectivity of judging and be consistent with most other judging systems of sports similar to freestyle disc (i.e., cannot be objectively measured). In addition, it would add to the professionalism and seriousness of our sport.

While this sounds logical and progressive, the reality in Freestyle Disc tournaments currently is that tournament staff are challenged with identifying 3 judges per category (for 9 total judges), much less 5 judges for AI, 5 for Diff, and 3 for Execution (totalling 13). Still the judging committee would like to note this as a progressive idea within the new judging manual as a **desired procedure, leaving it up to tournament directors to decide if enough qualified manpower is available at their tournament to realise it**.

To reduce the required manpower a bit, we discussed that it would be acceptable to have 2 judges for Execution only, one of which could be the head judge at major tournaments (so 12 instead of 10 judges would be required per pool), since the head judge in the current system has no active duties during routines. Execution is pretty objective and leaves little room for interpretation, so we don't really need 3 judges here (they are doing it in Footbag like this). If there are major Ex judging mistakes, they can easily be proven by video or witnesses and corrected afterwards. Of course having less judges for Execution doesn't mean that it should have less weight for the overall score as well, so the categories would have to be mathematically rebalanced.

## 6. Judging education

This was consistently and frequently mentioned to be a significant problem behind many of the judging system deficiencies we discussed. So **quantity and quality of trainings has to be increased, and the FPA Board will have to work out and implement new clinics based on the new judging manual**. Once this is done we think that players and tournament directors should be **incentivised to complete the clinics**. The FPA board has set up a task force to steer this process.