



Prague, March 8th 2023

WDSF AGM 2023 – Czech Motion No. 1: Stop AJS

Problem description

AJS in version 1.0 was introduced in WDSF competitions in 2009. The version AJS 2.0 was introduced in 2013 and the version AJS2.1 in 2015. The current version AJS 3.1 is in use since 2020.

The main goal was to create a new evaluation system that:

- a) should be based on ISU evaluation system, which was accepted by IOC,
- b) should improve the objectivity of the evaluation,
- c) should improve transparency,
- d) should improve the understanding of the evaluation in the IOC, the media and the public,
- e) should provide constructive feedback to couples and coaches.

Although the WDSF has devoted considerable effort to develop a new evaluation system, the stated goals have not been met, and its current form and use leads to the degradation of evaluation at top competitions.

In particular, the main problems are:

- 1) the possibility to influence the overall result even by one judge. It is enough to calculate the limits of the tolerance field and move within its range in accordance with your interest.
- 2) the incomparability of the evaluation at different competitions, which is due to the different settings of the judges, not the absolute performances of the couples. Couple marking often differs based on the other dancing couples, so in the presence of better couples you receive a lower marking, while in the presence of worse couples you receive higher marks. The differences are usually more than 1 point of absolute evaluation! "Negativity/positivity" and the tendency of the adjudicators are also not solved, so the evaluation is very depended on the composition of the panel of adjudicators. This leads to differences of almost 2 points!
- 3) accepting the lack of discernment of adjudicators, who can give the same rating to multiple performances. Ties should occur infrequently or only in a small number of couples, but in the results we see far more tie than match the evaluated performances.
- 4) the effort to create an evaluation interesting for viewers. In the past so-called open marking was used whith skating. The audience's misunderstanding of the difference in marking and the "psychological protection of couples" (so that they would not be stressed by a negative rating) were the reasons for its cancellation. But we started stressing the couples again and offered the viewers aggregate numbers to distract attention from the direct way to the result.



Therefore, we suggest:

1. to stop the AJS 3.1 at all competitions,
2. to return to the Revised final format using skating and supplement it with the video presentation of results for spectators at all competitions where the AJS 3.1 system was used,
3. to simplify and refine the PAS/IQ system even to lower levels and to create a unified description of the evaluation criteria for all competitions organised by WDSF and its national members, which will replace the earlier criteria systems associated with skating (ICAD, DTV).

Detailed explanation

To the main goals of AJS:

ad a) **ISU rules** are used by Program Components, Technical Element Scores and Deductions. AJS1.0 was designed this way, but only Program components remained in subsequent versions, which were further elaborated in the Adjudicator's Handbook as the PAS system. It defined 4 evaluation components for both disciplines (TQ/MM/PS/CP) further elaborated up to 12/3/4/6 subcomponents and used several IQ in each. Everything was based on solo dances, which is only used in a minority of final dances now.

Several experts participated in the creation of this material, and its imbalance remained a problem. Because there was an effort to define previously unnamed areas and principles and to capture the rapid development of dance sport, problems arose with the nomenclature and description of the newly captured principles. The development of this material stopped in 2017 and was only interpreted differently at adjudicator congresses. Despite all the shortcomings, this material can be evaluated positively. However, it needs to be balanced and further modernized. The biggest problems appear to be:

- imbalance of subcriteria,
- lack of clarity in the definitions and in the used IQs,
- the possibility of the same marking, which is caused by an insufficiently fine scale or an insufficient ability to distinguish or alibi of the adjudicators,
- unresolved lower levels,
- the efforts to create different scales according to the expected (desired) level of different categories,
- the effort to define everything up to the highest level blocked further development and excluded the aesthetic and artistic part of our dance, which were very subjective and difficult to grasp categories, but which are the most interesting thing about our sport (X-factor)
- unresolved link to the evaluation system used in the elimination rounds (skating does not solve any criteria),
- most definitions are based on solo dances and their use in group dances, which is used more, is not resolved,
- to use this system for other and new disciplines.

ad b) **Improving the objectivity** of the assessment should have been achieved primarily:

- by ensuring the same conditions for evaluation for all adjudicators. All adjudicators should see the same thing, i.e. the whole dance from the same place - from this only the same reserved places for the adjudicators are observed and even this has not been done recently (just watch the judges so that did not see the published marking). However, by starting to use absolute judging for group dances, the basic advantage of the system, that everyone can see everything, was completely suppressed.
- by dividing the judges into groups for each component, which are constantly changing and the composition of the groups also changes - this proved to be very difficult for the judges to fulfill and gradually became simpler until the current division into only two groups evaluating two components (TQ+PS = 16 subcomponents, MM+CP = 9 subcomponents).
- by separating the direct determination of the ranking on the rating scale and the mathematical processing of the individual ratings, which reacts to the deviation from the median rating - the judges quickly learned to use this system, and even though several statistical methods were tried to eliminate "fake" ratings, it is visible that this system is very easy to abuse even by a single adjudicator.

In this area, it is necessary to focus on:

- division of components and subcomponents so that some of the judges are not overloaded, others are not.
- searching for other mathematical methods that would be better able to determine the order from partial evaluations (statistical methods need large data sets, otherwise their results are inconclusive).
- performing evaluation analyzes not only with the aim of punishing the adjudicators, but also, for example, to determine the fineness of the scales, tolerance ranges, statistical parameters of the adjudicators, etc. Focus in particular on evaluations that are inconclusive, as they may show the adjudicators' insufficient discrimination ability.

ad c) **The transparency** of evaluations does not need to be addressed, because it is solved by publishing complete results and everyone can find them on the Internet in a clear form. The publication of intermediate results in the finals on the central screens is rather interesting for viewers and it can be a motivational element for couples. In this respect, it is a positive element, but it often delays the progress of the competition. The same result was achieved in the past by the so-called public evaluation of the final.

ad d) **Understanding** of evaluation in the IOC, media and public - nothing has changed in this area, because media outputs and PR matter has not improved in any way. It doesn't really matter if we explain 5 criteria in the ICAD system or 4 components in AJS. The public rather expects a correct evaluation of the difficulty in solo performances, as is the case with other aesthetic sports. The paradox is that figure skaters publish points for technical value during the performance and then just add the program components to the result, and we only use difficult-to-explain program components and want to publish them in a comprehensible way. So all the viewers understand is the value of the numbers they see on the screen and from which they can deduce the order. Link to performed performances, especially in the group dance, is missing.

ad e) Providing **constructive feedback** to couples and coaches is rather just a wish and a myth of AJS. Combining the two ratings into one number (even if we count it twice) is more of a scam on couples and coaches than any useful information for them. Couples receive this information more often at

training camps, where the PAS system is used for match analysis.

General conclusion

The AJS system was introduced with good intentions, but very amateurishly. Most of the experts were politically excluded from the necessary discussions, and further adjustments were prompted by commercial rather than professional sporting requirements. This diluted the basic principles of absolute evaluation, and the competitions are returning to their old ways.

The AJS system is used in less than 5% of competitions, yet it is used by all adjudicators congresses. No attention is paid to the rating system used in more than 95% of the competitions, for which the WDSF doesn't even have a proper name.

Unfortunately, there was no mention of how coaches and couples should react to the AJS system, how the strategy and preparation for the competition need to be changed.

We haven't made it to the Olympics yet, so we have time to stop, rethink everything, discuss and consider:

- The basic principle of our competition – groups or solo, or both? For solo we must have AJS, to compare is better in groups. Wouldn't the so-called Revised format, which solves this problem, be enough? And why is the already approved battle system not used?
- The "pentathlon/decathlon" system (collecting points for the overall result) or skating (each dance has its final order)?
- Creation of uniform systems of evaluation components for all rounds of competitions and for all levels.
- **Cardinal question - how important is music?** The first criterium according to ICAD, DTV (>95% of competitions) or only 25% of performance according to AJS (<5% of competitions)?