

Report of the FEI Dressage Task Force

October 16 2009

Table of Contents:

1. Introduction

2. Members of the Task Force Group

3. Tasks given by the FEI Bureau

4. The welfare of the horse

5. The structure of the Dressage Committee

6. Judging

7. Review format of Olympic Games

8. Qualification for Championships & Games

9. Series in Dressage

10. Miscellaneous

Annex 1. What members of the Judges Supervisory Panel could correct

Annex 2. Report from the Judging Systems Trial in Aachen 7-9th September 2009

Annex 3. Education System for Dressage Judges

Annex 4. Possible judging system to standardise the judging of the degree of difficulty in Freestyle tests

1. INTRODUCTION by Frank Kemperman, Chairman of the FEI Dressage Task Force

On the occasion of its autumn meeting in November 2008, the FEI Bureau decided to appoint a Dressage Task Force (DTF) to look at specific issues that are listed below.

Other issues and pending matters were not discussed by the DTF, as the group was not installed as a normal Technical Committee. Only on request of the FEI staff could the DTF be asked to advise on other matters. The issue of 'Welfare of the horse' was not originally included in the tasks for the DTF. However, since the DTF feels that this is a very important issue permission was given for it to be included in the final report. A further request was granted to include the item entitled "series in the dressage sport".

The DTF convened 7 days of physical meetings and held conference calls almost every 14 days. At its first meeting the DTF agreed to ensure that after every meeting or as often as would be required, a Summary Report would be circulated to all the stakeholders. Apart from National Federations (NFs) the FEI defined stakeholder groups to include all trainers, riders, judges and organisers regardless of whether they were members, or not, of their relevant recognised Clubs. As the DTF was made up of representatives of these different stakeholder groups it was agreed that each person would be the key point of contact for the group they represented on the DTF.

In this report the DTF presents its recommendations for the future development of the dressage sport to the FEI Bureau. Addressing the popularity of the dressage sport, the DTF was unanimous in highlighting the need to increase transparency and attractiveness for spectators, media and sponsors.

In many countries Dressage is already a growing sport. It is hoped that the proposals in this report will help to strengthen the development of dressage and take it to a wider audience.

I would like to thank the staff in the FEI headquarters, and all who have contributed to this report for their input and expertise, but special thanks must go to my colleagues, and fellow members of the DTF, not only for the endless hours of fair and intense discussions, but also for their unfailing commitment and responsibility each one has dedicated to this project.

In presenting the DTF's deliberations, advice and recommendations it is our humble wish that, in accordance with the FEI Bureau's original brief, we have fulfilled the tasks bestowed upon us.

It will be for the new Dressage Committee to determine the timeframe for the implementation of the proposals however, the DTF emphasises that it is a high priority to implement many of these issues as quickly as possible.

2. MEMBERS OF THE TASK FORCE GROUP

Name	Country	Representing
• Frank Kemperman (Chair)	(NED)	Organizers
• Richard Davison	(GBR)	Riders
• Robert Dover	(USA)	Trainers
• Alain Francqueville	(FRA)	Chefs d'Equipe
• Elisabeth Max-Theurer	(AUT)	Owners
• Katrina Wuest	(GER)	Judges

The above mentioned were chosen by the FEI Bureau and come from different stakeholder groups. They were not functioning as representatives of the associations of the stakeholders. Each person has given their personal opinion into the conversations of the DTF.

3. TASKS GIVEN BY FEI TO TASK FORCE GROUP

Following brief was given to the DTF by the Bureau:

1. Review the issues surrounding the very significant area of the training and development, assessment and selection of judges for major championships and Olympic Games; review as part of this development of randomised / computerised judge selection process;
2. The fitness for purpose of the method of judging Dressage competitions needs thorough review – both in terms of the number of judges, their positioning and the judging process;
3. Following significant feedback from NFs, it is clear that the decision to move from four to three riders in a team is not universally accepted as the best for the sport and this therefore needs reviewing;
4. The system for qualification for Championships, and the receipt of Certificates of Capability for Championships, World Cups, and the Olympic Games;
5. Review the consultation process within Dressage and how it affects the structure of the committee going forward to ensure maximum involvement by the key stakeholders, both internal to the sport and external within the greater sporting/Olympic/Paralympic environment.

4. THE WELFARE OF THE HORSE

The welfare of the horse is the key issue for the future of the sport. Sport with an animal will only be accepted by the public as long as we can convince the world that we have respect for the horse and we avoid any form of abuse or the use of prohibited substances. The discussions about doping and abuse to horses in the last few months have shown the world that everybody involved must do everything possible to avoid such negative messages.

The proposed rules for Equine Anti-Doping and Medication Control are the first step to reach this goal. It is absolutely necessary to have better rules, but it is also very important that everybody involved in the sport understands that they have to behave according to the FEI Code of Conduct for the Welfare of the Horse. All involved should understand that we cannot afford more scandals. Athletes, Owners, Grooms, Vets, Organizers and Officials should make every effort to ensure there is always utmost respect for our partner the horse.

We can discuss for hours and hours about rules, judging systems and other “important” issues, but we should first show the world that equestrian sport is based on a fair partnership between human and horse.

5. STRUCTURE OF THE FEI DRESSAGE COMMITTEE

One of the most urgent issues to discuss for the DTF was the structure of the new FEI Dressage Committee (DC), as the new Committee has to be nominated in November 2009. The DTF had been asked to recommend the structure of the future DC to the Bureau Meeting at the end of March 2009 to give all stakeholders the opportunity to give their comments and to make it possible to implement the new structure at the General Assembly 2009, and to select a new Dressage Committee.

After consultation with stakeholders, a proposal in which all stakeholders should have a seat in the Technical Committee was made. The DTF fully realised that the proposal was not in line with the actual statutes and regulations. The DTF was surprised that the Bureau members as well as the Chair of the Nominations Committee were of the opinion that the new DC should be in line with the existing regulations as opposed to looking at representation on the Technical Committees in a different way in line with the structure of the DTF itself.

On April 29th the DTF made the following new recommendation to the Bureau regarding the make up of the Dressage Committee, which is in line with the regulations.

The Riders and Organizers representative will be nominated in line with Statute 29.10 *"Candidates to any Standing Committee can be proposed by such Standing Committee, the Bureau or National federations, or, in the case of representatives of Athletes and Organizers, by the relevant FEI recognized international Association."*

The remaining representatives of the Committee will be nominated by NF's in line with Statute 32.2. *"Technical Committees shall include individuals with skills and experience in the following, as appropriate: competition, judging, event organization, training, and course designing."*

The DTF recognizes the need to ensure that the Committee has a good balance of skills and that it is made up of members who are of high quality with strong experience and knowledge in the sport of Dressage. It is also recognized that a good geographical spread is important in choosing Committee members. Experts from FEI HQ should advise the Committee in areas such as Development, Commercial and Communications when required. Also other external experts can be invited when required.

The Committee members should all be of different nationalities i.e. maximum one per NF. Gender and geographical spread is recommended, however having the right people is seen as more important. The issue of how to involve developing countries more needs to be further looked into.

6. JUDGING

The most emotional issue in dressage is judging. The DTF would like to remark that the sport should have respect for officials who are prepared to judge long competitions for a small fee then to receive sometimes negative criticism after the competition.

Not only in dressage but in many other sports there are a lot of discussions about the judging principles and processes. For that reason the DTF invited experts from other sports in order to compare different judging systems. Several ideas came up and thanks to the financial support of the FEI the DTF organized a trial event to test several judging systems and ideas.

At the Judging Systems Trial, which was organized on September 7-9 in Aachen, a team of very experienced judges of the highest level and some lower level judges were invited. After the trial the outcome of the Test was discussed and the opinion of the judges involved was important for the decision making process of the DTF.

6a. Training and Education

Good judging starts with good education!

Consequently, a new Judges Education System for all levels of judges has been worked out to guarantee their best possible education all over the world. This system, still in draft form, was presented and scrutinised. It could come into force January 2010 at the earliest.

The Education System includes:

- the Education Strategy
- the Entry Requirements for Judges' Qualification Courses
- the Course Organisation
- a standardised Examination Procedure (incl. consequences in the case of failure)
- the Criteria for FEI Judges to remain listed
- the Removal from the list and a possible way back
- the Codex to be signed by all Judges

It is important that all FEI judges are trained regularly and re-examined when necessary. "Refresher Seminars" without exam, "Sit-Ins" with experienced judges, and "Shadow Judging" shall provide worldwide training opportunities. In addition, these procedures are a good tool for the FEI to assess their judges as well as a globally applicable possibility for the up-and-coming national judges to participate in the FEI system.

In the future, important issues such as media training or musical education (for the evaluation of Freestyle Competitions) will be part of the course syllabus.

There is a strong need for better definitions regarding the Course Directors, how they are selected, trained and evaluated. Please refer to Annex 3 for further details.

Proposal:

Good education is the basis for good judging. The Education System will be renewed and globally improved by regular training, assessment, and re-examination of FEI judges on all levels.

6b. Evaluation and assessment

The following resources will be used for an evaluation and assessment of the quality of judging:

1. Judges Supervisory Panel
2. Foreign Judges
3. Consistency statistics
4. Informal resources, such as experience from events, feedback from stakeholders, officials, Dressage Committee and others.

As in normal business and in many other sports it is often necessary to organise a good system to evaluate judges. It is proposed to create a Judges Supervisory Panel (JSP).

This Panel of observers should consist of one core group of 3-4 independent, experienced, well respected judges and/or trainers/riders. The names of the JSP members must be given well in advance every year. Every effort must be made to avoid conflicts of interest. Additional members can be added, both to cover the quantitative and geographical needs. The JSP members will be proposed by the Dressage Committee to the FEI HQ who will make the final decision. Those who are not judges should be required to take actively and successfully part in a crash course in judging. Please refer to Annex 1 for further details.

One of the main tasks of the JSP should be to evaluate the judges at events. Other tasks of the JSP will be described in the following paragraphs. The JSP's function should be further developed in the future.

Members of this Panel would, as a start, go to Games and Continental Championships, which are at Grand Prix level, and the World Cup Final. They should observe all classes. The JSP will have to make a confidential evaluation report to the FEI / DC after each event. As soon as possible also normal CDIs should be visited by JSP members.

Several judges' evaluation techniques were discussed in detail, including use of a judge's evaluation form for analysis and feedback from appropriate judges, trainers and riders and official debriefing meetings. A proposed Judges evaluation form should cover topics such as personal ranking of competitors, use of the scale of marks, quality of remarks, objectiveness and independence and openness to discussion.

David Stickland has developed a system to analyse the consistency in judging. The results of this objective comparison can never be used in isolation but it is a very good additional tool to assist the JSP in evaluating the judges' performance.

On events where the JSP is not present the Foreign Judge should evaluate, organise and chair the debriefing. It is important that these meetings take place after the first principal competition, like the Grand Prix, rather than at the end of the event. This would allow the judges to reflect on their performance and possibly improve their results during the remainder of the show.

It is suggested that in order to take this forwards a further pilot project is run. Initially there should be a group of 3-4 people operating as the JSP in order to see how the evaluation system functions. The costs for the JSP should be paid by the FEI and the OC. It was suggested that the JSP members could also act as Appeal Committee members to help reduce overall costs.

Proposal:

Evaluation of judges is of great importance and the system has to be improved. A Judges Supervisory Panel should be installed and start to function for each Games and Continental Championship at Grand Prix level and the World Cup Final. The JSP is responsible for the evaluation of the judges during the event. The JSP will report to and advise the FEI about the quality of the judges. This proposal should be implemented as quickly as possible and as a high priority.

As the JSP will evaluate the judges at the events, it should also be the task of the JSP to advise the DC and FEI HQ regarding the promotion or demotion of the judges. For that reason it is necessary to have an overall Panel, consisting of 3 persons. This panel will collect all information, analyse the results from all CDIs worldwide and advise the DC and FEI HQ.

Judges should be awarded the status (3, 4 or 5 *) for a limited period of 3 years. After this period the status of the judge will be reviewed by the JSP. It is possible to promote or demote a judge before the end of this period, in the case the judge is functioning extremely well or badly. For every judge a "logbook" will be required with all details regarding events, evaluation, education etc.

Proposal:

Assessment and evaluation of judges will be the task of the JSP. A core panel will advise the DC and FEI HQ about the status of the judges. The judges status will be for a period of 3 years only. It is possible to change the status at any time. A logbook with all relevant information will be created.

6c. Selection of judges

The existing selection process for Championships and Olympic Games was discussed. The DTF was informed about the Quality Criteria for Nomination of Officials and the appointment procedure that had been used by the FEI Staff and Dressage Committee in previous years. In the past the FEI Staff consulted the DC members, Organising Committees and NFs of the respective Championships as well as the International Dressage Riders Club (IDRC) and International Dressage Trainers Club (IDTC) and asked for proposals, and compiled the names given in the responses. The DC Chairman and FEI HQ together then decided who the judges would be.

The quality criteria for nomination were discussed in detail by the DTF. It is a sensitive and also a political issue. The DTF is of the opinion that generally speaking the best judges should officiate at Championships and Olympic Games. With a good evaluation procedure and advice by the JSP the DC should advise the Dressage Director about the judges for the highlight events.

The DTF proposes the following procedure for the selection of judges for Championships, and other main-events.:

1. The JSP makes a proposal to the DC.
2. The DC advises the FEI HQ

Regarding the appointment of judges it was discussed to publish first a long list of judges several months in advance of the main events and a short (definite) list a predetermined number of months before the concerned event. With this system the names would be confidential until a few months before the event. As the DTF did not come to a conclusion the next DC should look at this proposal again.

For CDIs 3*/4*/5* the OC chooses the judges based on the existing criteria in the rules. The Foreign Judge is only appointed by the FEI for World Cup qualifiers in the Western European League and Games/Championships. In all events it should be preferably an O-judge, if available, who acts as the Foreign Judge. Organisers will be advised not to invite the same judges every year as the rotation of judges is preferable for the sport.

Proposal:

The best judges should officiate at Championships and Olympic Games. Only the JSP should advise the DC and FEI HQ regarding the names for the judges for the highlight events, based on the judges' record and objective criteria

6d. Judging method

An extensive judging systems trial was conducted to give statistical material for the conclusions below. This was held in Aachen 7-9 September. Results from 39 Grand Prix rides and 16 Intermediaire I rides created the statistical material for the conclusions below. Please refer to Annex 2 for further details.

6d.1. Number of judges

At this moment there are 7 judges present at selected Games and Championships from whom 5 are judging each competition. In order to reduce the influence of each judge's marks from 20% to 14.2%, it was proposed that all 7 judges are used in all competitions at these events.

At the test event the results with 7 judges and 5 judges for reference were compared. On this occasion the concept of dropping the highest and lowest results from the final score and per movement was analysed.

The statistics showed that there was no significant effect on final scores or ranks, either when all 7 judges' scores counted or by dropping scores. The introduction of the JSP will secure that technical mistakes or numerical mistakes by a judge will be corrected.

A number of analysis were conducted regarding removing the highest and lowest scores – this was conducted using actual results from past shows, and it was found that in these cases only small differences in the results occurred. It was therefore felt that this would not add a benefit to the sport.

Proposal:

It is recommended to use 7 judges at Olympic Games and Continental Championships on Grand Prix level, with all 7 judges counting to objectively reduce the influence of each single judge's marks.

6d.2. Use of half marks

The use of half marks was also tested. By using the training system of the judges actively, using half marks will give more accurate and consistent results. In the test, the ranking of the riders was not changed. However, it will give more flexibility for the judges, and the feedback to the riders in their score sheets will be more accurate. Today, it can be confusing for the rider to have a full mark difference for two movements which are more or less identical. The test gave the opportunity to compare with the marks the judges would have given with only whole marks. The half marks were used in both directions; up and down, but with a tendency to go up. The DTF is of the opinion that half marks will improve the percentages.

Proposal:

It is recommended to implement half marks for all levels of shows and classes. No change will be made in the Young Horses classes, where 0.1 decimals are already in use.

6d.3. Dividing of tasks

The 3 systems that were tested:

- a. Dividing the Judges tasks for normal competitions by having one set of judges looking at the marks for the movements and one set focusing on 7 Summary Marks giving an overall picture as in Young Horse tests
- b. Dividing the judges tasks for normal competitions in a significantly different way from the current system with each judge judging 2 different aspects of the performance
- c. For Freestyle competitions dividing the tasks between technical marks and artistic marks

In many other sports the tasks of the judges are divided. For that reason a lot of work was spent to create methods to judge the tests with dividing the tasks for the judges. This was tested both in the standards tests (Grand Prix) and in the Freestyle. The DTF does not recommend splitting the tasks in standard tests for the time being, as the collective marks are much more closely related to the movement marks than in Freestyle competitions.

Especially in the Freestyle to Music the dividing of tasks for technical and artistic marks worked extremely well and the system gave the judges more time to judge and concentrate on the test, especially the artistic marks. There is an uneven number of judges in every judges panel. To secure the most accurate scores, the highest number of judges will judge the artistic part. (4 judges giving artistic marks, 3 judges giving technical marks when 7 judges and 3 and 2 when 5 judges).

At the same time a new system developed by the DTF was tested which standardises the judging of the degree of difficulty. All systems which transform parts of the judging from subjective to objective reduce the uncertainty and question of personal opinions. Under this suggested system riders have to provide information in advance about the choreography of their tests. The difficulties above the compulsory minimum must be indicated including the possible bonus. The first impression is that this system could be very interesting and should be worked out further, but it needs to be simplified from today's one. Please refer to Annex 4 for further details.

Proposal:

It is recommended to divide the tasks between technical execution and artistic performance in Freestyle. It is also recommended to change the Collective Marks for all tests.

6d.4. Positions of the judges

For many years the positions of the judges around the arena has remained unchanged. Having all judges viewing from the same angle was tested. 5 judges were positioned at the short side (C) and in the same test 5 judges were positioned at the long side. Having the two judges normally described as "H" and "M" judges moved to sit at the long side at H and M was also tested.

The result of having all judges on the same side showed a slight decrease of differences between the judges, but too small to be significantly meaningful. The DTF finds it important that the judges cover as many angles as possible to have a correct total result.

All tests done with various positions showed that the position is of less importance than assumed. At almost all events the judges' huts block the view for the spectators. It is possible to be more flexible than today and hence help the organizers to find better solutions to improve visibility. This should be a matter for the next DC.

Proposal:

The DTF recommends establishing guidelines for judges' positions. In the case that there are 2 additional judges they will normally sit at either side of A. In situations where there are reasons to change placing from the normal, the FEI can make exceptions within the limits of these guidelines. All possible variations should secure a good total view of the arena for the judges' panel.

6e. Anonymity

In some sports the results per judge are not known by the public. The DTF agrees that the judging panel's total result is the correct one. The advantage of such a system is that there is less discussion about an individual person's judging. However the DTF finds that this would be contradictory to the clear goal of more transparency in the sport. The judges are also positive about having their own scores identified. As the judges are representing the FEI and not the individual NF, the nationality should be removed from the lists.

7. REVIEW FORMAT AND PARTICIPATION FOR THE OLYMPIC GAMES

At the Olympic Games, the highlight for every sport, dressage has the chance to present itself in the best possible way. For that reason the Dressage Task Force is proposing a more attractive and transparent format and a change to the participation structure:

7a. Participation

For the 2008 Olympic Games the number of team members was reduced from four to three in order to have greater universality. Equestrian sports currently have 42 NFs competing at the Olympics, ranking 18th (middle) of all sports in a ranking of number of participating countries.

Number of dressage participants:

	Athens 2004	Hong Kong 2008
Athletes	52	47
Teams	10	11 (12 qualified)
Individuals/NF	12 /7	10/9
Total NF	17	22

There were five more NFs in Hong Kong 2008 than in Athens 2004. A percentage of 64% was required to be eligible, however out of 47 starters in Hong Kong only 28 scored over 64%. The question is whether universality or quality of the sport is more important. According to the IOC qualification rules the best should participate but the same principles state that universality (geographical spread) is important in order to give developing countries a chance. The Dressage Task Force fully understand the importance to develop the sport and make it more global, but feel that it is important that the world's best athletes and horses are participating at the Olympic Games.

The Dressage Task Force feel that the former reserve athlete/horse should participate as individuals. Reserve riders should not be required to travel long distances without being able to compete. For the NFs qualified with a team it must be possible to bring a team of 3 riders and 1 individual rider. The 4th rider should start as an individual and at the same time act as a substitute for the team. Under the Task Force proposals NFs would choose their team at the latest 1 hour after the horse-inspection. Composite teams should be able to have a fourth athlete competing individually as long as they are qualified from the Olympic ranking list.

This would mean that there are, as in Hong Kong 2008, no drop scores in the team competition. In order to make the sport more understandable this system is very much preferred by TV and visitors without detailed knowledge of the sport.

To keep the total number to the allocated 50, only one team from Group F/G would qualify from a FEI approved 2011 Asian Pacific Dressage Championship and one team from Group D/E would qualify from the 2011 Pan American Games. Under the new proposals the following would be the situation for Hong Kong compared to those who actually qualified for Hong Kong:

	Hong Kong qualified	Hong Kong new proposal
Athletes	50	50
Teams	12	10 (8 direct, 2 composite)
Individuals/NF	14/11	12/11
Total NF	23	21

All numbers above are based on who qualified, not who finally participated

Proposal:
Under the new proposals the following NFs would have qualified Teams for Hong Kong: GER, NED, USA, SWE, SUI, GBR, AUS, CAN, all with 4 members (BRA and JPN would have lost the team qualification). The key changes have been highlighted:

DRESSAGE TEAM – QUALIFICATION SYSTEM

			Reserve list (R)
1.	The 3 best placed teams from the 2010 World Equestrian Games, Kentucky (USA)	3	
2.	The 3 best placed teams from the 2011 European Championship, excluding teams qualified above	3	
3.	The one best placed team from the Olympic Group F/G from a FEI approved 2011 Asian Pacific Dressage Championship, excluding the teams qualified above	1 (Previously 2)	
4.	The one best team from the Olympic Group D/E of the 2011 Pan American Games, excluding the teams qualified above	1 (Previously 2)	
TOTAL		8 teams (32 riders)	

Proposal:
“Composite” qualified teams: 3 or 4 individuals qualified from the same nation from the Olympic rankings to make up a team. Note in Hong Kong the 9th and 10th Teams were composite – Denmark and Spain however under the existing system and under the proposed system they would remain Teams of 3.

DRESSAGE INDIVIDUAL – QUALIFICATION SYSTEM

The following qualification places for Individuals may only be allocated to NOCs which have NOT qualified teams “directly”. Each qualification place will be for one athlete and one horse.

			Reserve list (R)
1	The host nation (GBR) will be qualified with 1 individual, if not otherwise qualified	(GBR) 1	
2.	Each NOC of the competitor placed first in the FEI Olympic Riders Ranking – Dressage, 1 May 2012, in each of the following regions or regional groupings will become qualified to enter 1 individual: A North West Europe B South West Europe C Central & Eastern Europe, Central Asia D North America E Central & South America F Africa & Middle East G South East Asia, Oceania	7	
3.	The completion of quota, either to fill up the spaces available or following cancellation of team or individual entries by a nation will be achieved by taking the athletes in their order of classification in the FEI Olympic Riders Ranking – Dressage, of 1 May 2012 to reach a total of 50 riders.	10	
4.	In accordance with the specifications noted under the heading ‘Tripartite Commission (IOC-ANOC-FEI) Invitation places’, Invitation places may be distributed to NOCs if confirmed by the Tripartite Commission		
	TOTAL	18 riders	

Proposal:

The DTF believes that the best combinations should compete at the Olympic Games. The former reserve rider should be activated as an individual rider and at the same function as substitution rider. The qualified teams could bring 3 athletes and 1 individual athlete.

7b. Competition format

The format of the competition has remained unchanged for many years and the DTF discussed the various options to create a more attractive programme for all involved especially the media and spectators. It is proposed to change the format but to maintain the same number of competition days. .

The schedule would be as follows:

- Day 1 & 2: Competition 1: first team-competition, open for all athletes (teams of 3, plus individuals). If a horse or athlete has to withdraw for medical reasons the fourth athlete will be a substitute and count in the team.
- Day 3 Competition 2: final team competition, open for 36 athletes (3 team members) (ie the 8 best teams and the 12 best individuals after competition 1. The individuals could come from the same nations as the teams if placed among the 12 best individuals) The individual could be used as a substitute for veterinary or medical reasons which disqualify the combination from participating further. However their score from Competition 1 will not be counted towards the final team score.
- Day 4 Competition 3: Individual competition, open for the best 18 athletes from Competitions 1 & 2 (max 3 per NF).

The main arguments for this format would be:

- Create the same possibilities as in the other equestrian disciplines for more athletes to compete more than once
- Enhance universality as more participants would be able to ride more than once
- Minimal impact on the length of the competition. The number of days would be the same as in Hong Kong 2008. The first competition would be unchanged. The second competition would be a short technical test thus the total time would be almost identical, and could even be shorter than today. The third competition would be only 30 minutes longer.

A more exciting starting order for Competition 2 is necessary:

- a) individual riders ranked 7-12 in reverse order from the result of Comp 1.
- b) the teams ranked 5th-8th after Comp 1,
- c) individual riders ranked 1-6 in reverse order from result of Comp 1.
- d) teams ranked 1st-4th, in reverse order. The last 12 riders, which will only take 2 hours, will normally decide the medals.

The first competition should be a technical test to give the judges the possibility to see if the combination shows the general principles of dressage. The second competition would be a short technical test. The third competition would be the Freestyle to Music as at the moment. Competitions 1 and 2 would decide the Team ranking and also serve as qualifiers for the Individual Final. The 18 best would qualify for the only individual medal, the Freestyle to Music.

This would mean that, as in other equestrian disciplines, dressage has only one individual winner. Only Competition 3 would decide the individual medals.

This proposal is based on the proposal that was forwarded to NFs in the spring and which at that time was accepted by the majority of NFs as a good step forwards.

Proposal:

The second Competition should be the Final for the Team Ranking and the format should be changed in the way that the last 12 competitors decide the Team medals. As currently is the case, only one individual medal should be rewarded and be decided in the Freestyle to Music test.

7c. Minimum eligibility criteria for participation in Olympic Games

Today the requirement is that all participating athletes/horses must have obtained at least 64% by two 5* judges of other nationalities than themselves. To enhance the neutrality, it should be added that also the average result of all judges in the qualifying competitions should be at least 64%.

Proposal:

The 64% criterion has to be obtained BOTH by a 5* judge AND as an average from all judges in the competition.

8. MINIMUM ELIGIBILITY CRITERIA FOR PARTICIPATION IN CHAMPIONSHIPS

Currently the qualification criteria for Championships are the same as for the Olympic Games - 64% in a GP by two 5* judges in two different CDIs. The same change as above should be introduced for Qualification for Championships. This should be introduced for Qualification for Continental Championships in 2011.

Proposal:

Keep the qualification system for championships as it is today, only adding that the 64% criterion has to be obtained BOTH by a 5* judge AND as an average from all judges in the competition. The same criterion is used for the Olympic Games and Championships.

9. SERIES IN DRESSAGE

To deliver dressage to its rightful position in the world of sports we have to create a product that is interesting for the public, the media, the athletes and especially for the sponsors. We need a product that is the highest, most compelling representation of the sport.

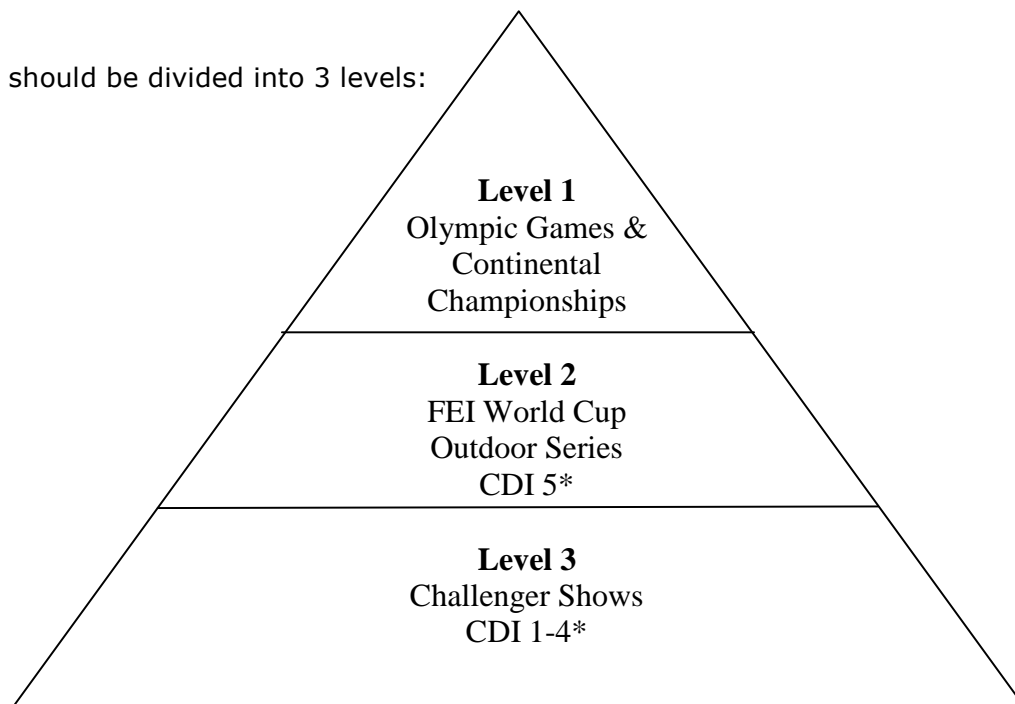
At the moment the dressage calendar consists of many international events. The Olympic Games, FEI World or Continental Championships are, or should be, the yearly highlights. During the (European) winter season the FEI World Cup is organised. The image of the World Cup is good, but could still be improved. Depending on the activities of the concerned Organizing Committees, all other events in the calendar are organised more or less successfully. At the same time some new products are launched, like for example the World Dressage Masters series or a Top 10 Final.

The situation in Show Jumping, with an overkill of series, has shown us that the FEI has to take the leading role. This means the FEI should come up with a strategy for series in Dressage. For this reason the Task Force has taken the initiative to make a proposal for a strategy.

STRUCTURE

The first step should be to structure the sport. In 2008 ca. 280 international dressage events were organised world wide. Dressage events need to be further developed and improved to make dressage a more popular sport. The fact is that the outside world does not understand what the real highlights are in our sport. Especially for the (non equestrian) media the situation is not understandable. The decision to start a categorisation of events, with a star system is very good, but not enough.

The events should be divided into 3 levels:



Level 1

The Olympic Games is the best possible platform in the world to show that dressage is an attractive and interesting sport; it belongs to the best sports in the world. In the report of the DTF we have made several proposals to make the dressage sport more attractive at the Olympic Games.

The creation of the FEI World Equestrian Games has shown us that the equestrian sport in general can get much more attention from media, sponsors and spectators. Here we can only hope that the FEI will find good organisers and that the chosen organiser will give dressage the best possible platform. The new format proposed by the DTF for the Olympic Games could also make the Championships much more attractive.

The Continental Championships are not always very successful. In the past it has happened that this Championship was only a party for the "in crowd" or did not have the media- or spectator attention as it could have. With the concept to organise Dressage at the same time and venue together with other discipline(s), the sport could get more media attention. It is difficult to find good organisers all over the continent and for that reason the idea to look for a new concept for Championships could be interesting.

Level 2

a) FEI World Cup

The FEI World Cup exists since 1985 and has become a well known product. The series is organised over 4 leagues and also has a development role for the sport.

The 4 leagues are:

- West Europe (10 events)
- North America (12 events)
- East Europe (9 events)
- Pacific (4 events)

The Finals really are a highlight.

What could be done to create a better FEI World Cup series?

- More quality regarding the competitors. Not all top riders are participating in the series. The series should be made more attractive through more prize-money.
- Extra World Ranking List points should be awarded.
- More quality regarding the events. Currently shows are organised at venues that have traditionally been part of the series. However, in the mean time other venues have developed to be more successful events of good quality. For that reason the selection criteria of events should be changed.
- The number of events in West Europe should be reduced to maximum 8. A maximum of one event per NF is desired.
- A sponsor should be found for the series.

b) Outdoor Series

To avoid a situation with several different kinds of series it is necessary to create one good product. Dressage riders and especially dressage horses do not compete in many competitions in one year. For that reason it will only be possible to create one series for the top level horses. This means that a series for the Top could only consist of a few events. Nation's Cup (CDIO) events are a traditional and highly valued product. At the moment some CDIO's are organised, but a link between the CDIO's is missing.

A new series of high level events organising a CDIO should be started.

- Maximum 4 events should be part of the series.
- It is important to find dates fitting in the calendar, date protection is necessary.
- The prize-money for the 3 Grand Prix competitions should be min. € 150.000.
- A series ranking for the teams and for the individual ranking has to be made.
- The proposed "Olympic Format" should be used for the series.
- If possible an extra bonus for the overall winner should be awarded.
- Extra World Ranking List points should be awarded.
- The top riders should commit to the series and sign a contract to confirm their participation.
- The best teams should participate (selection according to World Ranking List).
- There should be a promotion / relegation rule for the last placed team(s).

As soon as the FEI agrees to the general idea, a detailed draft should be made and events should be selected. With a good concept and commitment of the top riders it must be possible to find a sponsor for the series.

25 and Under

To help young riders to make the step from the young-riders level to Grand Prix level, an extra series for riders in the 25 and under age group should be created and included in the new CDIO series. At the same events extra competitions for Nations teams of 2 riders should be scheduled. The participating teams should be from the same countries as the seniors and these riders and officials should coach the 25 and under riders. Also for these 25 and under riders a team and individual ranking will be made.

c) CDI-5*

Other CDI 5 stars events will also be categorised under Level 2. These events are not part of series but will have a place in the calendar.

d) Top 10 final

The idea to create a Top 10 Final for dressage will not be easy to realise. One reason is that in dressage not the riders but the combinations are ranked. This means that riders have more horses placed in the Top 10. The FEI should assure that the rights of a "Top 10 Final" are with the FEI. An event like this can only be organized if the Top 10 are really participating, otherwise it is only bad image for the sport.

Level 3

It is of great importance that many CDI events in the category 1 - 4* are organised. These events are important for the sport, but are not the highlights for the media.

10. MISCELLANEOUS

Further recommendations by the DTF:

- The DTF asks the DC to look at the number of tests in each level and each age group. The DTF is of the opinion that more tests at each level should be created
- The words used to describe each mark should be reviewed
- The naming of the classes should be reviewed
- The height, construction and materials for the judges boxes should be looked at especially in relation to improving the visibility of the arena for the audience
- The prize giving ceremonies should be reviewed in order to establish the correct balance between the safety of the horse and rider and the attractiveness to the media

Categorisation of events

For some years now events are categorised only according to the level of prize money. It is absolutely necessary that the events are also classified according to other criteria such as:

- Quality of footing
- Quality of competition arena
- Number and quality of training areas
- Accommodation for horses and riders
- Interest from spectators and media
- Efficiency of organisation

The FEI foreign judge and the foreign rider should make a report after each event and the shows should be categorised according to the results. New events should start on a max. 3 star level (independent of the prize-money) and in the 2nd year they could get a higher classification. In the last months an event evaluation system for show-jumping has been discussed and a concept is drafted. The work was done by the Education Department of the FEI together with a small group of organizers in which the chairman of the DTF was involved. This evaluation system could easily be modified and introduced for dressage.

Modern techniques

To bring dressage closer to a wider public it is necessary to look at the possibilities regarding the use of modern technical instruments. The use of open scoring is a very well accepted system to involve the spectators with the competition - this system should be obligatory for the higher level events. Together with experts from TV the use of graphic systems, spider cams and other new technical possibilities should be explored.

During the last year some proposals were made to use technical systems for judging. In some cases the systems did not look usable and/or the costs of the system made it unattractive. The DTF is of the opinion that every new technical development should be looked at and be evaluated for its usefulness in dressage sport.

Annex I: Catalogue of mistakes to be corrected by the JSP during the course of a test

Conditions for the correction:

- I. A judge is clearly higher than his colleagues (minimum 2 points)
- II. The marks of the other judges should be 5 or below

Mistakes that could be corrected:

1. Numerical mistakes

- a. in series of flying changes
- b. in canter zig-zags

2. Clear and definite technical mistakes

- a. single changes late behind
- b. clear changes of leg in pirouettes
- c. every technical mistake that leads to a mark of 5 or below, **unnoticed by a judge** (e.g. unnoticed breaking into canter in extended trot, passage, half pass, short moment of jogging in walk)

The correction mark does not have to be the average of the other judges' marks. It is up to the JSP to decide on the correction mark.

Initial Report on the Judging Systems Trials held at Aachen, September 6-7 2009

Dr. David Stickland

Executive Summary

1. Test1a: Seven judges.

- Removing scores based on final scores of course improved the consistency between the remaining judges, but had no significant effect on final scores or ranks. Removing them movement-by-movement changed final scores in a few cases by up to 0.6%.
- It may not be cost effective to invite 7 judges but to use only 5 scores.
- Removing outlying movement scores, has little effect, and the goals would be better achieved by targeting truly anomalous scores using for example the Judge Supervisory Panel

2. Test 1b: Component scoring system

- Remarkably good agreement was achieved given the radical nature of the system. The system may be useful today in judge training, further analysis and development of the details would be required before it could be actually used in competition.

3. Test 2: Half-Points

- The judges used half-points about 25% of the time, with roughly as many upward as downward corrections to the standard system. They were able to use them consistently. There is the good reason to suppose that with practice the use of half-points would improve precision

4. Test 3: Separating Collective and Technical scores for Grand Prix

- While the rankings were more or less the same, the individual judges using the new collective scores showed a wide variation in final scores, possibly due to the small total number of notes that they can give. It would not seem to be a wise choice to base so much on such a few marks, the likelihood is that this would distort the final results more than improve their precision.

5. Test 4: Separating Artistic and Technical scoring in the Freestyle

- This test was popular with the judges as it allowed the "artistic" judges to concentrate on those aspects of the performance. The results were encouraging and the few ranking differences were ascribed to genuine effects.

6. Test 5: Long-side and Short-side differences

- There was clear evidence of consistency between judges on the same side of the arena and their inconsistency with judges on the adjacent side. This supports the (current) wide distribution of judges around the arena to be sure that the performance can be properly assessed from all angles

Test 1a

The description of this test is given below:

1. 7 judges judging, Grand Prix test positions E,H,C,M,B + both sides of A), otherwise today's system (using scale 0-10, giving marks for all movements and collective marks)

This will let us test:

- a. 5 judges as today for reference (by selecting the results of E,H,C,M,B judges)
- b. 7 judges all counting
- c. 7 judges, 5 counting, removing highest and lowest score (final score)

Nine riders participated to this test, and the results are summarized in Table 1

Rider	E	H	C	M	B	F	K	7 Judge Average	7 Judge Rank	Closest 5	7-5 Difference	7 Judge Max-Min	5 Judge Max-Min
4	65.11	66.17	67.87	67.66	62.98	63.19	62.98	65.14	1	65.02	0.12	4.89	4.47
2	62.77	63.40	64.68	64.04	64.04	59.36	62.98	63.04	2	63.45	-0.41	5.32	1.28
6	60.43	60.00	63.40	63.62	61.91	61.70	65.32	62.34	3	62.21	0.13	5.32	3.19
9	61.06	61.49	61.06	62.55	64.47	59.15	65.32	62.16	4	62.13	0.03	6.17	3.40
1	61.49	61.91	61.91	64.04	61.91	60.21	61.91	61.91	5	61.83	0.09	3.83	0.00
3	59.57	61.91	60.21	62.77	61.06	58.09	58.94	60.36	6	60.34	0.02	4.68	2.98
5	58.30	55.74	60.43	59.36	62.77	60.64	60.64	59.70	7	59.87	-0.18	7.02	2.34
8	60.00	56.81	59.57	58.30	59.36	56.38	58.72	58.45	8	58.55	-0.10	3.62	2.77
7	57.02	58.94	57.45	57.45	57.23	56.38	61.06	57.93	9	57.62	0.32	4.68	1.91
Averages											0.0	5.1	2.5

Table 1 Test1a, 7 judges using standard system. The highest and lowest judges scores are highlighted in red and blue. Removing these two judges does not change the ranking at all, however it reduces the average range of marks by a factor of two from about 5 to 2.5.

In this first treatment we compare the scores from the 7 judges with that obtained by removing the two judges with the highest and lowest scores. While the consistency of the 5 remaining judges is inevitably better when the two extremes are removed, there is in this case no change of ranking nor is there a large change of final score.

In Table 2 we show a similar result, but this time removing (arbitrarily) the two "new" judges (Those at F,K on the A-shortsides)

Rider	E	H	C	M	B	F	K	7 Judge Average	7 Judge Rank	Standard 5 Judges	7-5 Difference	7 Judge Max-Min	5 Judge Max-Min
4	65.11	66.17	67.87	67.66	62.98	63.19	62.98	65.14	1	65.96	-0.82	4.89	4.89
2	62.77	63.40	64.68	64.04	64.04	59.36	62.98	63.04	2	63.79	-0.75	5.32	1.91
6	60.43	60.00	63.40	63.62	61.91	61.70	65.32	62.34	3	61.87	0.47	5.32	3.62
9	61.06	61.49	61.06	62.55	64.47	59.15	65.32	62.16	4	62.13	0.03	6.17	3.40
1	61.49	61.91	61.91	64.04	61.91	60.21	61.91	61.91	5	62.26	-0.34	3.83	2.55
3	59.57	61.91	60.21	62.77	61.06	58.09	58.94	60.36	6	61.11	-0.74	4.68	3.19
5	58.30	55.74	60.43	59.36	62.77	60.64	60.64	59.70	7	59.32	0.38	7.02	7.02
8	60.00	56.81	59.57	58.30	59.36	56.38	58.72	58.45	8	58.81	-0.36	3.62	3.19
7	57.02	58.94	57.45	57.45	57.23	56.38	61.06	57.93	9	57.62	0.32	4.68	1.91
Averages											-0.2	5.1	3.5

Table 2 Test1a, this time removing (arbitrarily) the scores from the two new judge positions (F,K). In this case there are small changes of rank that are highlighted in yellow. The total range of scores is also reduced by a smaller factor from 5.1 to 3.5.

The statistical sample is small (only 9 riders) but the test shows that removing outlying final scores achieves (inevitably) the desired reduction in range of scores, but makes no significant real difference in the final scores or the ranking. Probably the rider is happier to see a range of scores that are closer together, but there is no real effect on the final result.

Undoubtedly having 7 judges **does** reduce the negative, or positive, effects of any single one of them on the final score, and so in the absence of economic factors might be desirable. But it is probably not cost-effective to introduce 7 judges instead of 5 and then remove the scores of two of them.

We can also study the effect of removing the top and bottom scores from each figure, rather than from the final scores.

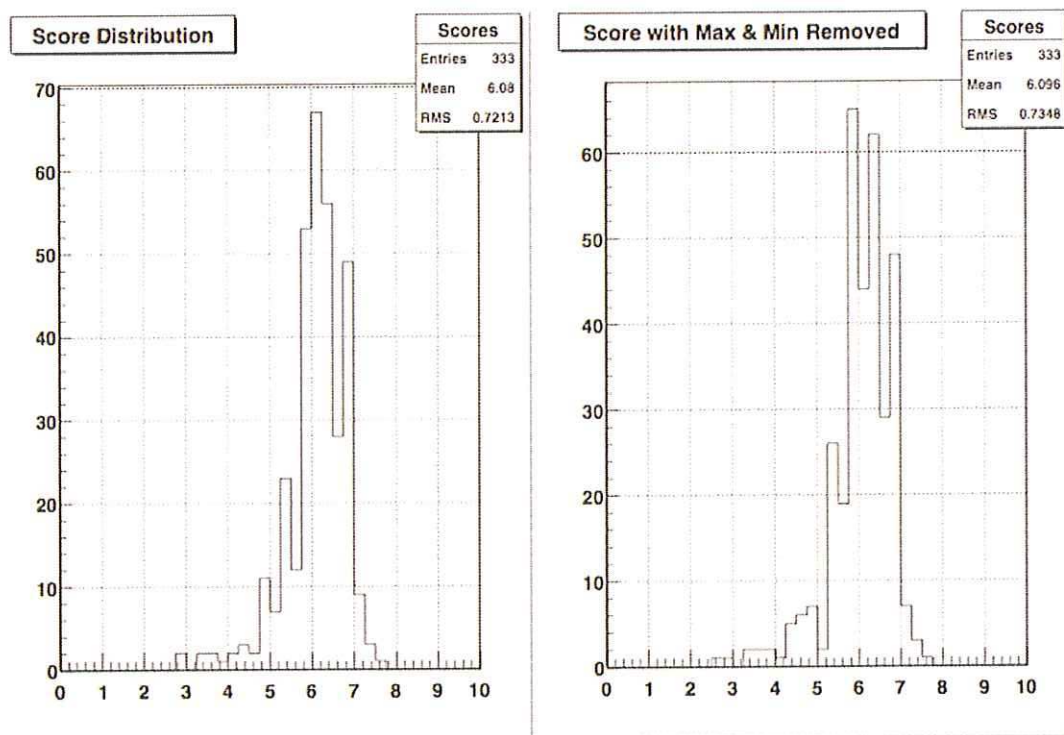


Figure 3 Score distributions, in the left hand plot we show the average score assigned by the seven judges for each figure; in the right hand lot we show the average for the remaining 5 judges after the highest and lowest scores are removed for each movement. Note that the Means and RMS (That is the spread) are virtually unchanged, and indeed the entire distributions are essentially identical

In Figure 3 we show the distribution of scores given in the first case by the average of 7 judges, and in the second case by the average of the 5 judges remaining after the top and bottom score have been removed for each figure.

We see that averaged over all movements in the 9 tests, the effect of removing top and bottom scores is very small. There is no significant effect on the mean score per figure. The effect on the RMS spread is that it changes from 0.721 to 0.735, this RMS is measured with a precision of about 0.004, so this is also only a marginally significant observation.

If we restrict the analysis to the first 5 judges (E,H,C,M&B), removing again the top and bottom scores for each figure, the means of these distributions are essentially unchanged, but the RMS spread changes from 0.703 to 0.713, again with marginal significance in this difference

In each case, removing judges from the final average slightly increases the RMS spread in the average score for each figure, this is quite probably just because the average is now an average of a smaller number of scores, and thus the average itself is less well measured.

In Table 4 we show the effect on the final scores of removing the highest and lowest scores on a movement by movement basis. There are three changes of rank but the average final score change is only 0.1%.

Rider	7 Judges average	7 Judges rank	Removing extreme scores Figure by Figure	New Ranking	Score Difference	Rank Difference
4	65.14	1	65.32	1	0.18	0
2	63.04	2	63.23	2	0.19	0
6	62.34	3	62.21	5	-0.13	-2
9	62.16	4	62.47	3	0.31	1
1	61.91	5	62.47	3	0.56	2
3	60.36	6	60.30	6	-0.06	0
5	59.70	7	59.57	7	-0.13	0
8	58.45	8	58.68	8	0.23	0
7	57.93	9	58.00	9	0.07	0

Table 4. Final scores after removal of the highest and lowest notes on a movement by movement basis. Rank changes are indicated in yellow.

Aachen Test Event								All 7 judges figure scores	Figure scores with top/bottom removed	Difference
Rider	E	H	C	M	B	F	K			
GRAND PRIX (2009)										
1	6	6	6	6	6	6	5	5.9	6.0	0.1
2	6	6	6	6	7	7	7	6.4	6.4	0.0
3	6	6	6	6	6	6	6	12.0	12.0	0.0
4	6	7	7	6	6	7	7	13.1	13.2	0.1
5	6	6	6	6	6	6	6	6.0	6.0	0.0
6	6	6	7	7	7	7	7	6.7	6.8	0.1
7	6	6	5	6	6	6	6	5.9	6.0	0.1
8	5	6	5	6	6	4	5	5.3	5.4	0.1
9	6	6	6	6	6	5	6	5.9	6.0	0.1
10	6	6	6	6	6	5	6	5.9	6.0	0.1
11	7	7	6	7	7	7	7	13.7	14.0	0.3
12	7	6	7	7	7	6	7	13.4	13.6	0.2
13	5	6	6	7	5	6	6	5.9	5.8	-0.1
14	6	6	6	6	6	5	5	5.7	5.8	0.1
15	5	6	6	6	6	5	6	5.7	5.8	0.1
16	6	6	6	6	6	6	6	6.0	6.0	0.0
17	6	6	6	6	5	4	6	5.6	5.8	0.2
18	7	7	7	7	7	7	7	7.0	7.0	0.0
19	7	7	7	7	7	7	7	7.0	7.0	0.0
20	6	7	7	8	7	7	7	7.0	7.0	0.0
21	7	7	7	8	6	7	7	7.0	7.0	0.0
22	6	5	6	6	6	6	6	11.7	12.0	0.3
23	5	6	6	6	5	4	5	10.6	10.8	0.2
24	5	6	5	5	5	4	4	9.7	9.6	-0.1
25	7	6	7	7	7	6	6	6.6	6.6	0.0
26	7	6	5	6	7	6	6	12.3	12.4	0.1
27	6	6	7	7	6	7	7	6.6	6.6	0.0
28	6	7	6	7	7	7	7	6.7	6.8	0.1
29	6	6	6	7	6	6	6	6.1	6.0	-0.1
30	6	6	6	6	5	6	6	5.9	6.0	0.1
31	7	7	6	7	6	6	6	6.4	6.4	0.0
32	5	6	6	6	6	6	6	5.9	6.0	0.1
33	7	7	7	7	7	8	7	7.1	7.0	-0.1
34	7	6	7	7	6	7	7	6.7	6.8	0.1
35	6	6	6	6	6	6	6	6.0	6.0	0.0
36	6	6	6	6	6	6	6	12.0	12.0	0.0
37	7	6	7	7	7	7	7	13.7	14.0	0.3

Table 5. The effect of removal of top and bottom scores on a particular rider. See text for details

Finally, we can examine in detail the effect of extreme score removal on Rider1, where the final score was changed upwards by 0.56%. This is shown in Table 5. This table also neatly illustrates that often many judges give the same score, consider line 1 (6,6,6,6,6,5) where the maximum score is 6, the minimum is 5, but of course we cannot tell which score of 6 we remove. In red we highlight the maximum scores, and in blue the minimum, only the scores that are neither a maximum nor a minimum are in white. We see that the final shift of 0.56% is comprised of many small shifts the largest of which is +0.3 points. Nothing stands out as being very strange. In fact what stands out is how close the judges were to each other in almost every movement.

I conclude that reducing the number of judges contributing to a final score is statistically not a sensible thing to do. If however, one judge were scoring "very poorly" then the correct approach would be to identify that judge and remove all of their scores, (or possibly just those that are wildly divergent from the other judges – however this second approach seems to be philosophically wrong as it is equivalent to just keeping the scores that have no effect on the final result, thus also rendering this judges scoring as being "irrelevant"). This begs the question of how one determines that a judge is indeed judging "very poorly".

Test 1b

The description of this test is given below:

1. *7 judges judging, Grand Prix test positions E,H,C,M,B + both sides of A), otherwise today's system (using scale 0-10, giving marks for all movements and collective marks)*
 - b. *4 judges, judging various elements of training scale,*

In this test 4 judges used a radically different judging system, being asked to evaluate not the entirety of the movement, but its separate components (Impulsion, Rhythm, Contact, Suppleness, Precision, Submission, Collection and Straightness). This clearly required considerable effort from the judges, and some questions arose about for example the definition of straightness in a lateral movement. Nevertheless the judges put maximum effort into this and came up with a remarkably consistent ranking. Ignoring all coefficients, and indeed any collective-marks, the final percentage scores were also remarkably close as shown in Table 6.

	Impulsion	Rk	Rhythm	Rk	Contact	Rk	Suppleness	Rk	Precision	Rk	Submission	Rk	Collection	Rk	Straightness	Rk	Total	Rk	7 Judges	Rk
Rider 4	68	2	65	3	72	1	68	1	82	1	71	1	61	1	62	2	68.63	1	65.14	1
Rider 2	63	4	67	1	57	6	57	8	73	2	68	2	60	3	57	7	62.88	3	63.04	2
Rider 6	60	7	60	8	70	2	66	3	66	6	62	6	58	6	58	4	62.70	4	62.34	3
Rider 9	62	5	57	9	68	3	68	2	71	4	64	4	60	2	62	1	63.98	2	62.16	4
Rider 1	59	9	62	5	59	5	65	5	69	5	65	3	58	4	60	3	62.06	6	61.91	5
Rider 3	69	1	67	2	53	9	58	7	73	3	63	5	58	4	58	4	62.50	5	60.36	6
Rider 5	60	7	62	5	66	4	65	4	65	7	59	8	56	8	56	8	61.07	7	59.70	7
Rider 8	61	6	61	7	56	7	58	6	65	7	59	7	56	8	58	4	59.14	8	58.45	8
Rider 7	66	3	64	4	54	8	53	9	63	9	55	9	58	7	56	9	58.49	9	57.93	9

Table 6. Component judging system. The average score for a rider in each component and their rank in that component is shown, as are a simple total score and by comparison the total score assigned by 7 judges using the standard system

In the table we highlight in yellow an interesting set of scores, where one rider/horse pair is ranked 1st and 2nd in Impulsion and Rhythm, but 9th and 7th in Contact and Suppleness. By comparison in the standard scoring system this rider received average collective marks of 7, 7, 5, 6.5 (Paces, Impulsion, Submission, Aids), the low mark for submission is probably reflected in the low component scores for Contact and Suppleness.

Further analysis will be required to extract pace-by-pace and/or movement-by-movement correlations between the component scoring and the standard system. Participants to the test expressed the feeling that this decomposition could be very useful in judge-training situations to encourage the judges to clearly identify the separate components that will eventually be combined to give a movement score. Clearly many details remain to be clarified before such a system could indeed be used in actual competition, but the results obtained here were promising and further analysis will be performed.

Test 2

The description of this test is given below:

2. *5 judges judging, Grand Prix test, positions E,H,C,M,B, giving half marks (using scale 0-10, giving marks for all movements and collective marks), and 5 judges judging as today's system*

This will let us test:

- a. *How often do the judges give half marks?*
- b. *Are the scores higher or lower?*
- c. *How is the distribution? More 6.5 instead of 6, more 7.5 instead of 8?*

In Table 7 we show the final results calculated using the two systems.

	Whole-points	Rank	Half-points	Rank	Diff
Rider2_6	67.70	1	68.17	1	0.47
Rider2_1	63.06	2	65.11	2	2.04
Rider2_3	62.72	3	64.04	3	1.32
Rider2_4	60.64	4	60.11	4	-0.53
Rider2_5	59.19	5	59.89	5	0.70
Rider2_7	59.06	6	57.60	6	-1.47
Rider2_2	56.98	7	57.55	7	0.57

Table 7. Comparison of final scores using the standard system with those obtained allowing half-points

There are no changes of rank, however the half-point judges were on average 0.5% higher than the standard system judges. The half-point judges were asked to note whether their choice of a half-point score was an upwards or a downwards effect compared to the score they would have given otherwise. Not all judges actually noted this, of those that did they stated that in 137 cases they went up with the half-point, and in 101 cases they went down. There is a suggestion in the scores in Table 7 that the tendency was actually to increase the scores for the higher ranked riders and to decrease them for the lower ranked riders. Anything that stretches the scale used (for valid reasons) would be an advantage in improving the accuracy of ranking, particularly for the riders with the "typical" score of around 63%-68%.

In Figure 8 we show the distribution of scores actually given by the standard system judges and by the half-points judges. The distributions are already quite similar and it is reasonable to expect that with more experience the judges will manage to fill the half-point scores even more uniformly so that the final score distribution becomes smoother and more precise separation of performances will be possible.

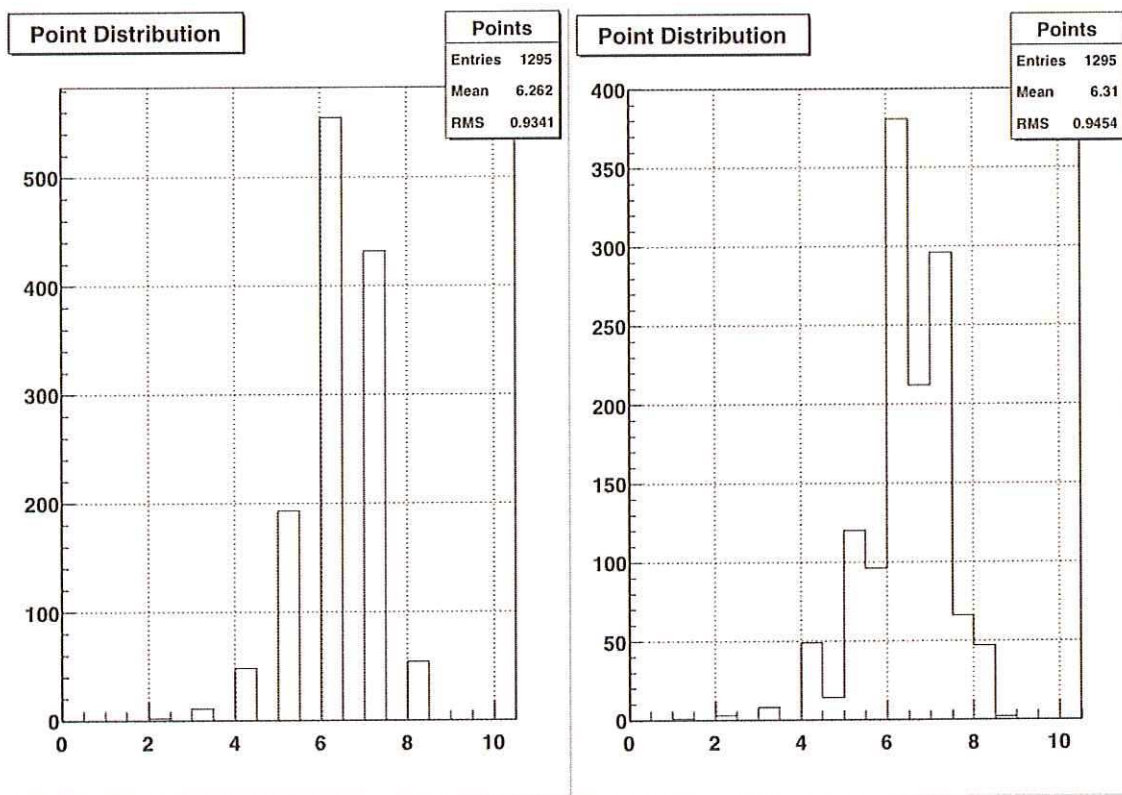


Figure 8. Points assigned in Test2, in the left hand plot the points given by the standard system judges, and in the right hand plot those given by the judges using half-points. The distributions are indeed very similar

The judges using half points assigned them 390 times in 8 riders (Out of a total of 1480 total judgements), they assigned 357 to the Technical marks and 33 to the collective marks, this is

	Half Points	Judgements	Percentage
Overall	390	1480	26%
Technical	357	1320	27%
Collective	33	160	21%

Table 9. Use of half-points. 26% of all judgements made use of the possibility of assigning half-points

summarized in Table 9.

The judges used half-points about 25% of the time. The constancy of the ranking and the distributions of points awarded support the hypothesis that judges are indeed well able to assign half-points consistently. The full analysis is not yet completed, but the observation is that very frequently 3 or more judges adopted the same half-point for a given figure, again suggesting they are already able to use such a tool precisely.

Test 3

The description of this test is given below:

3. *Dividing of tasks, standard tests Grand Prix test)*

*5 judges giving marks for movements (and collective marks, as today, for reference),
5 judges giving collective marks according to a newly designed testsheet.*

This will let us test:

- a. *Will the total marks be higher/lower than today?*
- b. *Will the variation between average marks for movement and average marks for collective marks be higher/lower than today?*

The newly designed test sheet contained marks for Trot, Walk, Canter, Submission, Quality of execution of movements, Rider's position and seat and General impression.

In table 10 we show the comparison of the results obtained by the standard judging, by the be collective marks, and for comparison, by the old collective marks.

	Standard	New Collective	Diff	Standard Collective	Diff
Rider3_3	66.30	70.0	3.7	69.0	1.0
Rider3_2	64.78	70.8	6.0	67.0	3.8
Rider3_1	62.61	68.6	6.0	65.0	3.6
Rider3_6	62.09	66.8	4.7	63.3	3.5
Rider3_11	62.04	67.1	5.1	64.7	2.4
Rider3_5	61.91	64.2	2.3	63.3	0.9
Rider3_8	59.61	62.4	2.8	60.3	2.1
Rider3_7	58.91	62.2	3.3	62.7	-0.5
Rider3_10	58.70	61.1	2.4	59.0	2.1
Rider3_4	57.43	59.5	2.1	60.0	-0.5
Rider3_9	53.48	55.0	1.5	53.3	1.7

Table 10. Results from Test3. A comparison of the scores obtained by standard judging and by judges using only a newly designed collective test-sheet. The judges with the new system were on average 3.6% above the standard system. However, even the old collective marking is also 1.8% above the standard score (Technical+Collective)

The judges using only the newly designed test sheets scored on average 3.6% higher than those using the standard system. However, as a comparison, we see that even in the standard system, the collective marks alone score on average 1.8% higher than the Technical+Collective marks. In yellow we note the rank changes.

We can examine in more detail the new system, in Table 11 we show the scores assigned by the 5 new-collective judges. The first thing that strikes you is the enormous range of scores assigned. For example Riders 11 and 5 have a score range of 11% and 15% respectively! Clearly the system is much more subjective and would be almost impossible to monitor or correct. Since the final scores are composed of only 7 separate marks, variations in one of them can introduce large variations in the final score.

New Collective Judges		H	C	M	B	Range	AVERAGE	Rank
Rider3 2	72.00	66.50	72.50	69.50	73.50	7	70.80	1
Rider3 3	72.50	70.00	66.00	70.00	71.50	7	70.00	2
Rider3 1	69.50	67.50	64.00	69.00	73.00	9	68.60	3
Rider3 11	62.00	65.00	68.50	67.00	73.00	11	67.10	4
Rider3 6	67.50	65.50	67.50	67.50	66.00	2	66.80	5
Rider3 5	56.50	63.50	64.50	71.50	65.00	15	64.20	6
Rider3 8	61.50	62.00	63.50	60.50	64.50	4	62.40	7
Rider3 7	62.00	62.00	60.50	62.50	64.00	4	62.20	8
Rider3 10	64.00	60.50	61.50	57.50	62.00	7	61.10	9
Rider3 4	60.50	62.00	60.50	60.50	54.00	8	59.50	10
Rider3 9	56.00	53.50	59.00	50.50	56.00	9	55.00	11

Table 10. Results from Test3- new collective marking system. the range of scores is very large for some riders, for example Rider 3_5 has scores ranging from 56.5% to 71.5%. Even the E-B judges have an almost 10% difference in score.

The use of this system alone is strongly disfavored by this test, the small number of score that are combined to make a final score, together with there subjectivity and difficulty to challenge, would not lead to a more objective scoring system.

Test 4

The description of this test is given below:

4. Dividing of tasks, GP Freestyle

4 judges judging the technical part as today from H, C, M and B, + giving half marks for final mark per movement. 3 judges, (B, between C and H, and between C and M) judge the artistic part with half points as normal and have the possibility to give more detailed remarks to the competitor.

This will let us test:

- a. *Will the total marks be higher/lower than today?*
- b. *Will the variation between average technical marks and average artistic marks be higher/lower than today?*

In Table 12 we show the comparison of the results obtained using the new Artistic test-sheet and the standard Technical test-sheet. The ranking is almost unchanged, the only significant change being that of Rider4_4. In discussions after the event the consensus was that this was a correct difference and that the musical harmony was really not at the same level as the Technical execution.

Artistic Only							
	E	C	M		Stdev	Average	Rank
Rider4 5	77.50	75.00	83.75		4.51	78.75	1
Rider4 8	70.00	77.50	78.75		4.73	75.42	2
Rider4 3	72.50	70.00	75.00		2.50	72.50	3
Rider4 1	66.25	73.75	72.50		4.02	70.83	4
Rider4 4	65.00	67.50	76.25		5.91	69.58	5
Rider4 7	66.25	67.50	65.00		1.25	66.25	6
Rider4 2	60.00	67.50	63.75		3.75	63.75	7
Rider4 6	57.50	66.25	62.50		4.39	62.08	8

Technical							
	K	H	B	F	Stdev	Average	Rank
Rider4 5	67.50	67.50	65.63	64.38	1.53	66.25	2
Rider4 8	66.04	65.83	61.67	67.50	2.51	65.26	3
Rider4 3	65.42	61.25	65.83	65.63	2.19	64.53	4
Rider4 1	66.67	62.92	62.71	61.25	2.31	63.39	5
Rider4 4	67.71	66.46	67.50	66.04	0.80	66.93	1
Rider4 7	61.04	60.21	58.13	61.67	1.54	60.26	8
Rider4 2	63.13	59.79	59.58	63.54	2.11	61.51	6
Rider4 6	61.25	60.63	60.00	62.29	0.98	61.04	7

Table 12. Results from Test4- Comparison of the marks given by the judges using a new Artistic-only test-sheet and those from the standard Technical test-sheet. We highlight in green the only significant rank change

The judges also expressed their appreciation of the opportunity to judge the Artistic performance separately, being able to concentrate on that part was considered a significant improvement. there was considerable and unresolved debate about whether some of the "Artistic" notes were in fact really "Technical" notes, and also as to how the artistic judges should best take into account Technical errors.

(The artistic scores given by the standard system are not currently available, but when they are it would be interesting to double-check if they are correlate well with those given by the Artistic-only judges.)

Test 5

The description of this test is given below:

5. *All at same side. 5 judges judging as today, but on short side + 5 judges judging as today, but on long side. 1 judge at opposite long side at M. Inter-I test*

This will let us test:

- a. *Comparison of marks for all movements assessed from the long as well as the short side.*
- b. *Are there any differences that should be taken into account or does it look the same from all positions?*

In Table 13 we address the simple question, Do the judges on the same side of the arena judge more closely together than in a normal arrangement? To compare the long-side, short-side arrangement with a standard arrangement, we have mixed two long-side and the three short-side judges at H,C,M to create also a "standard" arrangement of judges.

In Table 13 we show the range in scores (maximum - minimum) for each movement, for a standard judging arrangement, and for the short-side and long-side arrangements. In the three columns we show the average range, the percentage of times when the range is 2 or more (for example scores 6,7,7,7,8) and finally the percentage of times when the range is 3 or more). The standard arrangement indeed has a higher number of large deviations than either the short-side or the long-side arrangements. However the difference between the standard arrangement and the average of the new arrangements is typically only about 2 standard-deviations. This suggests that there is indeed an improvement in scoring consistency when the judges are on the same side, but it is not conclusive. (About 10% of the time such 2 standard-deviation effects will happen without indicating anything abnormal.)

	Average Range	2 or more	3 or more
Standard	1.0	20.7±2.3%	4.7±1.1%
Short-side	0.8	12.6±1.8%	2.2±0.7%
Long-Side	0.9	16.7±2.0%	0.7±0.4%

Table 13. Results from Test5- Are there fewer large deviations when the judges sit all on the same side? The data suggest that this is true, but they conclusions are not definitive due to the small sample size.

In Table 14 we show the Figure by Figure scores for those figures where the range of scores between the Long-side and the short-side is 3 or more.

When the average score difference between the two sides is large (say 1 or more) we are probably seeing effects due to a genuine difference in view between the two sides. For example the first "Changes" (row 4) shows an average (and consistent) scoring on the long side of "5", and an average and consistent scoring of "7" on the short-side. By contrast row 11 has consistent 6 scores on the long side, and 4 on the short. ***There are a number of examples like this, where it seems clear that the full view of the event is only obtained by having judges on at least a short and a long side of the arena.***

In Table 15 we show the final score and rank comparison between the long-side and the short-side judges. In Table 16 we explore in more detail the case of Rider5_14 where the final score difference is 2.9% (Normal Grand Prix judging has a precision per judge of about 1.5%, thus for a panel of 5 judges we expect a final precision of approximately $1.5/\sqrt{5}=0.7\%$, therefore a 2.9% difference in score for the two panels is about a 4 standard-deviation effect, which is significant from a statistical point of view.)

Row	Figure	L1	L2	L3	L4	L5	Long Ave	S1	S2	S3	S4	S5	Short Ave	Long-Short	Max(Long)-Min(Short)	Max(Short)-Min(Long)
1	Flying Change	6	7	6	7	6	6.4	7	3	7	7	6	6	0.4	4	1
2	Canter Pirouett	6	5	5	5	5	5.2	6	5	7	6	8	6.4	-1.2	1	3
3	Canter Pirouett	6	7	6	6	5	6	5	4	6	4	5	4.8	1.2	3	1
4	Changes	5	5	5	5	4	4.8	7	7	7	7	7	7	-2.2	-2	3
5	Changes	5	5	7	5	5	5.4	7	8	7	7	7	7.2	-1.8	0	3
6	Changes	5	5	7	5	5	5.4	5	4	4	4	5	4.4	1.0	3	0
7	Changes	7	8	7	6	7	7	5	5	5	5	5	5	2.0	3	-1
8	Changes	7	5	5	5	4	5.2	6	6	7	7	7	6.6	-1.4	1	3
9	Changes	5	6	7	6	7	6.2	6	6	7	6	4	5.8	0.4	3	2
10	Changes	4	4	4	5	4	4.2	5	7	4	3	4	4.6	-0.4	2	3
11	Changes	6	6	6	5	7	6	5	4	4	4	4	4.2	1.8	3	0
12	Changes	3	4	4	5	4	4	5	6	5	5	5	5.2	-1.2	0	3
13	Collected Cante	5	5	5	5	6	5.2	6	8	6	6	7	6.6	-1.4	0	3
14	Collected Walk	6	6	6	6	5	5.8	7	8	6	7	7	7	-1.2	0	3
15	Collected Walk	7	5	5	7	6	6	5	6	5	4	5	5	1.0	3	1
16	Extended Trot	4	7	7	5	5	5.6	7	7	7	7	7	7	-1.4	0	3
17	Extended Trot	7	7	8	6	6	6.8	6	7	6	5	5	5.8	1.0	3	1
18	Extended Trot	7	7	6	6	7	6.6	7	4	5	6	6	5.6	1.0	3	1
19	Extended Trot	6	7	7	6	7	6.6	6	4	4	6	7	5.4	1.2	3	1
20	Extended Walk	5	5	6	5	5	5.2	6	8	6	7	6	6.6	-1.4	0	3
21	Flying Change	5	5	4	6	6	5.2	6	7	7	6	7	6.6	-1.4	0	3
22	Flying Change	6	6	7	7	7	6.6	6	4	4	5	4	4.6	2.0	3	0
23	Flying Change	4	5	5	6	5	5	7	7	7	6	5	6.4	-1.4	1	3
24	Flying Change	6	6	7	6	7	6.4	7	5	5	4	4	5	1.4	3	1
25	Rein-Back	5	6	7	6	6	6	6	4	5	6	5	5.2	0.8	3	1
26	Shoulder In	7	7	7	7	7	7	6	5	6	7	4	5.6	1.4	3	0
27	Transition	6	6	7	6	7	6.4	6	4	6	6	5	5.4	1.0	3	0
28	Transition	6	4	5	5	5	5	6	7	7	7	7	6.8	-1.8	0	3
29	Trot	7	5	8	7	6	6.6	7	6	6	5	6	6	0.6	3	2
30	Trot	7	6	7	6	6	6.4	6	4	5	7	7	5.8	0.6	3	1

Table 14. Results from Test5. We show details of those figures where the range of scores between the short-side and long-side judges is 3 or more. The columns headed L1-L5 show the scores of the long-side judges (From A-end to C-end), the columns S1-S5 show the short-side judges, from H-M. The average Long and Short-side scores are shown in red, as is the difference (Long-Short). On the right we show the max difference between a Long and Short-side score.

	SHORT	Rk	LONG	Rk	Score Diff	Rank Diff
Rider5_1	68.63	1	66.68	3	-1.9	2
Rider5_10	67.16	2	68.00	1	0.8	-1
Rider5_7	66.32	3	65.16	4	-1.2	1
Rider5_14	65.05	4	67.95	2	2.9	-2
Rider5_13	64.63	5	64.58	5	-0.1	
Rider5_4	63.63	6	63.47	7	-0.2	1
Rider5_8	62.95	7	61.32	11	-1.6	4
Rider5_11	62.47	8	62.68	8	0.2	
Rider5_9	62.32	9	63.89	6	1.6	-3
Rider5_2	61.16	10	61.05	12	-0.1	2
Rider5_3	61.11	11	61.63	9	0.5	-2
Rider5_12	60.00	12	61.63	10	1.6	-2
Rider5_5	58.21	13	60.21	13	2.0	
Rider5_6	52.63	14	54.21	14	1.6	

Table 15. Results from Test5. The comparison of final rank and score for the case of 5 long-side judges and five short-side judges. In table 16 we investigate the case of Rider5_14, where the score difference is 2.9% between long-side and short-side judges

Table 16. Results from Test5. The case of Rider 5_14 where the short-side judges scored 2.9% less than the long side judges. In each case the internal agreement between the long-, or short-, side judges is good, so the difference seems to be clearly due to genuine perceived differences from the two viewpoints. On the right we note the effect of each movement on the difference in final scores, almost half of which comes from the 4 figures highlighted in red.

INTERMEDIATE I (2009)		Short Side					Long Side					Long-Short (effect on final score)
Rider5_14	Coef	H	HC	C	CM	M	K	V	E	S	H'	
Halt-immobility-salute	1	6	7	7	6	7	7	7	6	6	7	0.00
Extended trot	1	7	7	6	7	6	7	7	7	6	7	0.05
Shoulder-in right	1	7	7	7	7	7	7	8	7	7	8	0.11
Volte right 8m	1	6	7	7	7	7	7	6	6	7	7	-0.05
Half-pass right	2	7	7	7	7	7	7	8	7	7	8	0.21
Rein back 5 steps	1	5	5	5	6	5	6	6	6	6	5	0.16
Half-pass left	2	6	6	6	7	7	7	7	7	7	7	0.32
Volte left 8m	1	6	7	6	7	6	7	7	6	7	6	0.05
Shoulder-in left	1	7	6	7	7	6	7	8	7	6	7	0.11
Medium trot	1	7	7	7	7	7	7	7	7	7	7	0.00
Collected walk	2	6	6	6	8	6	7	7	7	7	7	0.32
Extended walk	2	7	6	6	6	6	6	6	7	7	6	0.11
Collected canter	1	7	7	6	7	7	6	7	7	7	7	0.00
3 half-passes 5m	2	7	7	6	7	6	7	7	7	7	7	0.21
Extended canter	1	7	7	7	6	6	7	7	6	7	7	0.05
Flying change of leg	1	7	7	7	7	7	6	7	6	7	7	-0.11
5 flying changes every 3rd	1	5	6	5	5	5	5	5	7	5	5	0.05
Pirouette left	2	5	6	6	6	6	7	7	7	6	5	0.32
Flying change of leg	1	7	7	7	7	7	6	7	7	7	7	-0.05
Pirouette right	2	6	5	6	6	6	6	7	6	6	6	0.21
Flying change of leg	1	7	7	7	7	7	6	7	7	6	7	-0.11
7 flying changes every 2nd	1	7	6	7	7	6	7	8	7	6	7	0.11
Extended trot	1	7	7	7	7	6	7	7	7	7	7	0.05
Transitions at C,M and K	1	7	7	7	7	7	6	7	7	7	7	-0.05
Halt-immobility-salute	1	7	7	6	7	7	7	6	7	7	7	0.00
Paces	1	7	6	7	7	7	7	8	7	7	7	0.11
Impulsion	1	7	7	6	7	6	7	7	6	7	8	0.11
Submission	2	7	6	6	6	6	8	7	7	7	6	0.42
Rider's position and seat	2	7	7	7	7	7	8	8	7	7	7	0.21
65.05%						67.95%						

Table 16 shows that the final score differences are largely due to consistent judging internally between the judges on one side of the arena. The 0.32% difference from the Left Half-pass may well reflect the much clearer view that the judges on the short side have of this movement (from H-X). While the Pirouettes are also performed much closer to the short-side judges and the collected walk is all but invisible to the long-side judges at H-M (on the long-sides). The collective mark difference comes 50% from the submission scores and 50% from the other collective marks, I invite judges to hypothesize why this might happen...

The clear consistency between judges on the same side and their clear inconsistency with the judges from the adjacent side seems to be strong evidence in support of the current system of having judges distributed as widely as possible around the arena to profit from different views in their overall assessment of the performance

Acknowledgments and further work

I am extremely grateful for the excellent assistance given to me by the SCG team at Aachen, in particular Marijn Dubbeldam and Jaco Weemstra onsite, and before the event by Arend Tolner. I thank Frank Kemperman and the Aachen crew for facilitating my work and the FEI for inviting me participate to these tests. I thank the many dressage riders, trainers and judges who have given me sound advice and helped me to obtain some insight into the judging process.

I welcome any and all suggestions for further analyses at email: davidstickland@me.com

Education System for Dressage Judges

Draft VII October 2009

1. Preliminary Remarks

Terminology and Abbreviations

- **Course Director (CD)**
Definition: CDs are responsible to deliver FEI QC, RC and RS. The group of CDs is also responsible for updating the existing FEI education material.
- **5* Judges Seminars**
To be defined
- **Dressage Committee (DC)**
- **Freestyle (FS)**
- **Grand Prix (GP)**
- **Grand Prix Special (GPS)**
- **Grand Prix Free Style (GPFS)**
- **FEI Headquarters (HQ)**
- **International Judges are divided into three categories:**
3* Judges (former International Candidate Judges)
4* Judges (former International Judges)
5* Judges (former Official International Judges)
- **Intermediate level 1 or 2 (Int. 1 or Int. 2)**
- **Judges Supervisory Panel (JSP)**
To be defined
- **Mentor**
A judge who supervises Site-ins. Normally it is a 4* judge, if possible 5 * judge (from another country than the candidate). The MJ is a approved by FEI
- **Qualification Course (QC)**
Definition: A course with exam to obtain the FEI Qualification as 3* or 4* judge.
The participation in a QC without taking the exam counts as participation in a RS.
- **Re-Qualification Course (RC)**
A course with exam to maintain the FEI Qualification as 3* or 4* judge. A RC is the same course as the QC. The participation in a RC without taking the exam counts as participation in a RS.

- **Refresher Seminar (RS)**
A seminar for the purpose of educating and updating actual FEI Judges. RS will be open for National Judges for education purposes in addition to the invited FEI Judges.
- **Shadow Judge (SJ)**
A judge who judges the class parallel to the official judges under supervision of the mentor judge.
- **Sit-In**
A judge who sits together with the official judge to learn about the judging and the process around it.

Education Strategy

The number of 3*, 4*, and 5* Judges for different regions and countries will be dependent on

- the number of international events
- the number of international riders
- the number of qualified candidates
- the role of an individual on national Level

The FEI will provide statistics on the actual balance between number of international events and number of FEI Judges per NF, Group and world-wide. A corresponding analysis will determine areas of various needs and the education strategy.

Nomination Process

A judge, who passed the examinations during a qualification course, can act as Judge on that particular level 1 month after the course. Corresponding certificates will be awarded during the closing ceremonies of the qualification course.

Judges Review

The Judges Supervisory Panel reviews all FEI listed Judges at least biannually. A report will be submitted to FEI Headquarters.

Information Resources on Quality of Judging

The following resources will be used for an evaluation of the quality of judging:

1. The Judges Supervisory Panel
2. The Foreign Judges Report
3. The Consistency Statistics
4. Informal Resources

These resources will be compiled and analysed within the FEI HQ also consulting the Dressage Committee.

Codex

To be signed by all Judges.

National Education...

Will be continued as separate project.

2. Diagram of the FEI Education System for Dressage Judges

To be developed

3. Description of the FEI Education System for Dressage Judges

3* Judge

Entry requirements for a 3* Judges Qualification Course

1. To have proven experience as competitor up to at least PSG Level, preferably GP. Proofs are provided either through results lists or to be confirmed by the NF.

Exceptions possible which have to be considered by the FEI (DC, HQ, JSP). In these cases compensation by: 2 more Sit-Ins and 2 more Shadow Judgings with an evaluation from the MJ graded "good" or better.
2. To be a certified National GP judge have judged a minimum of 9 national GP Level classes in the last three years prior to the application (proven by official results or NF written confirmation). Minimum 3 in last 12 month before exam and application.
3. To speak and write English Written confirmation by the NF and assessment of Mentor Judges during Sit Ins and Shadow Judging
4. Recommendation of respective NF Letter
5. Sit-Ins 3 Sit-Ins (Int. 2, minimum 1 GP/GPS, minimum 1 GPFS) with at least 2 out of 3 positive recommendations of different MJ (4*, whenever possible 5*). The MJs are to be approved by the FEI (HQ).
6. Shadow Judging 2 Shadow Judgings (GP, GPS, minimum 8 horses, 2 different shows) with 2 positive recommendations of different MJs. One of the MJs per Shadow Judging to be a 5* or 4* judge. The MJs are to be approved by the FEI (HQ) and forwarded to the applicant.

The Sit-Ins and the Shadow Judging must be conducted at a minimum of 2 CDIs.

Sit-Ins and Shadow Judging may be done at the same show but the Shadow Judging must be done before the Sit-In.

One of them should be a CDI 3* or above

Course Organisation 3* Judges Qualification Course

Number of days	3
Number of Course Directors	2
Minimum number of participants	10
Maximum number of participants for exams	15
Syllabus	To be developed
Debriefing	Average 10 Min per participant Depends
Awards Ceremonies	To be decided
Course Material	To be developed
CD Notes	To be developed

Course Examination 3* Judges Qualification Course

Practical Examination	Judging of GP class with min of 10 horses
Written Examination	Format and size to be decided
Oral Examination	Questions on - the practical judging - knowledge of the principles of riding / judging - FEI Dressage- and Freestyle Rules Exam duration, to be decided Number of questions, to be decided
Marking Criteria	To be developed
Final Evaluation	To be developed

Criteria to remain listed
3* Judges
Proof of activity within 3 years

Minimum of 12 international classes (except Young Horses classes) at a minimum of 4 CDI 2* above. 6 of the classes judged must be Intermediate II and higher.

Max. 3 classes on level Int. 2 or higher and max 3 classes on lower level can be replaced by either 1 positive Shadow Judging (per class to be judged) or the participation in FEI RS (per class to be judged).

MJs for the Shadow Judging, one of them being a 5* judge, are to be approved by the FEI (HQ).

Course with re-examination only when needed (taking into account assessment of DC, HQ, JSP, Consistency Statistics).

4* Judge
Entry requirements for a 4* Judges Qualification Course

1. Certified International 3* judge
For at least 2 years and have judged a minimum of 8 CDIs 2* and above including 12 classes on GP level
2. Sit-Ins
2 Sit-Ins (GP, GPS) with 2 positive recommendations of different MJ (4*, whenever possible 5*). The MJs are to be approved by the FEI (HQ).
3. Shadow Judging
2 Shadow Judging (GP, GPS, minimum 8 horses, 2 different shows) with 2 positive recommendations of different MJs. One of the MJs per Shadow Judging to be a 5* or 4* judge. The MJs are to be approved by the FEI (HQ) and forwarded to the applicant.

The Sit-Ins and the Shadow Judging must be conducted at a minimum of 2 CDIs.

Sit-Ins and Shadow Judging may be done at the same show but the Shadow Judging must be done before the Sit-In.

One of them should be a CDI 3* or above
4. NF notification
NF will be notified on the invitation and have the possibility to object. FEI (HQ) will consider the reasons for objection.
5. Application letter from the Judge to FEI
The Judge has to formally apply to FEI (HQ) for invitation to a 4* Judges Qualification Course, specifying within the application all necessary requirements.

Course Organisation 4* Judges Qualification Course

Number of days	3
Number of Course Directors	2
Minimum number of participants	10
Maximum number of participants for exams	15
Syllabus	To be developed
Debriefing	Average 10 Min per participant
Awards Ceremonies	To be decided
Course Material	To be developed
CD Notes	To be developed

Course Examination 4* Judges Qualification Course

Practical Examination	Judging of GP class with min of 10 horses
Written Examination	Format and size to be decided
Oral Examination	Questions on - the practical judging - knowledge of the principles of riding / judging - FEI Dressage- and Freestyle Rules Exam duration, to be decided Number of questions, to be decided
Marking Criteria	To be developed
Final Evaluation	To be developed

Criteria to remain listed
4* Judges
Proof of activity within 3 years

Minimum of 12 international Intermediate II and higher classes at a minimum of 8 CDI 2* and above.

Max of 3 classes can be replaced by either 1 positive Shadow Judging (per class to be judged) or the participation in a FEI RS (per class to be judged).

MJs for the Shadow Judging, one of them being a 5* judge, are to be approved by the FEI (HQ).

Course with re-examination only when needed (taking into account assessment of DC, HQ, JS Consistency Statistics average of minimum 3 years)

5* Judge
Promotion requirements for a 5* Judge

- | | |
|--|--|
| 1. Certified International 4* judge | For at least 3 years and have judged a minimum of 15 CDIs 3* and above including 30 classes on GP level |
| 2. To fulfil the conditions required to be a 5* Judge and the Guidelines for FEI 5* Judges | High quality of judging and high ethical standards. Able to give clinics, willing to develop the sport further, willing and be able to act as Foreign Judge, take over more responsibility |
| 3. Selected by FEI (HQ) | HQ proactively selects from the list of 4* Judges based on competence, knowledge, experience and appearance and after hearing the DC and JSP. |

Criteria to remain listed
5* Judges
Proof of activity within 3 years

Minimum of 12 internat. GP-Level classes at a minimum of 8 CDI 3* and above.

Attendance of 5* Judges Seminars every 2nd year and one RS every four years, if not acted as CD.

Available for further development of the education system.

Course with re-examination for 5* Judges only when needed (taking into account assessment of DC, HQ, JSP, Consistency Statistics)

4. Failure of Exams

If a judge first time fails an exam within a qualification or re-qualification course for 4* judges,

1. he/ she will have to re-do the exam within a period of 2 years
2. he/ she will have to do 3 Sit-Ins and 2 additional Shadow Judgings with recommendations of the MJ graded “good” or better as a result of Shadow Judging. MJs for the Shadow Judging, one of them being a 5* judge, are to be approved by the FEI (HQ).

If a second failure occurs, the Judge will be downgraded to the status of national judge. After 2 years the judge can try again by starting to fulfil the entry requirements for 3* judges.

5. Removal from and return to list

A judge can be taken off the FEI list for the following reasons, controlled by the FEI (HQ):

1. Inactivity, not fulfilling the requirements (considered by FEI HQ in a case-by-case evaluation) 1 successful FEI RS or positive 1 Shadow Judging to get back.
2. Poor quality of of judging (see Information Resources on Quality of Judging)) results in minimum 1 year off list
Poor quality to be specified
To get back on the list: to fulfil conditions of class again (Sit ins, Shadow Judging) and exam.
3. Proven to have acted seriously or repeatedly against guidelines, proven biased judging / Codex results in minimum 1 year off list
To get back on the list: to fulfil conditions of class again (Sit ins, Shadow Judging) and exam.

The JSP will decide in a case-by-case evaluation on the individual return conditions.

6. Course Directors

The FEI has to keep a sufficient number of course directors. The list shall be reviewed by HQ every 2nd year to secure the quality of the education. Course Directors will be selected, trained and monitored by the FEI HQ. A list of CDs is published on the FEI Website.

Degree of Difficulty in Freestyle Competitions

Can the assessment be standardized?

Criteria for the Degree of Difficulty

- Difficult movements
- Combinations of movements
- Difficult transitions
- Repetition of movements / parts of a movement
- Reins in 1 hand

Ext. paces not to be counted for the degree of difficulty

**Precondition for an increased Degree of Difficulty:
All movements, combinations, or transitions have to be well executed!**

The Neutral Level

Mark 6.0 = the „neutral“ level

showing **minimum** of requirements (acc. to test sheet)

- Showing each required movement only 1x
- No combinations of movements
- No difficult transitions
- Execution out of an easy approach

Minimal risk, execution technically correct

The Bonus-System, starting from 6.0

- All difficulties, combinations, difficult transitions, repetitions ... have to be defined by different **grades of difficulty**. These grades start from + 0.1 → + 0.3
- A group of experts has to define the bonus system and review it annually
- All difficulties, combinations, difficult transitions ... shown above the compulsory minimum will get
 - a) a bonus if executed **7 and above** or
 - b) +/- 0 (= neutral) if executed **6 → 6.9** or
 - c) a malus - corresponding to the bonus - if executed **below 6**

How can this system work in practice?

- Riders must hand in not only their music, but a **CD** with a „game plan“ of their Freestyle choreography (the day before)
(Computer applicable system to be developed by the FEI)
- Each choreography must be indicated in chronological order
- All difficulties, combinations, transitions, repetitions ... above the compulsory minimum must be indicated including the bonus that could be achieved for each movement
- A **printout** of the choreography has to be given to the judges the day before. This way, they can check and define the indicated degree of difficulty beforehand

Judges divide their tasks

Judges are divided into three groups:

- 1. Technical-Judges (→ technical execution)
- 2. Artistic-Judges (→ harmony, choreography and music)
- 3. Difficulty-Judges (→ degree of difficulty)
- Needed in addition: A Data operator

Task of the T-Judges/ Data Operator

- The T-Judges follow the choreography of each rider (printed out from the CD as individual testsheets and shown on the computer) and give their marks movement by movement as in a standard test
- The Data Operator receives the marks from the T-Judges and immediately sends the average of these marks to the D-Judges who can see how a movement is evaluated.

Task of the D-Judges

- 2 D-Judges sit together and evaluate jointly the degree of diff.
- a. following the average marks of the T-Judges
 - b. according to the bonus/malus-schema

They are equipped with a computer and an additional video system for replay

At the end they add up the bonus points and deduct the minus

If the rider changes his Freestyle, the D-judges have to react.

Leaving out a difficulty → no bonus points

Adding an extra → additional bonus points (if executed 7 +)

Task of the Data Operator

When the Freestyle performance is finished, the Data Operator adds up the marks of

- a) the T-Judges (for the technical movements) and divide them by ??? to get the final % = **1 mark**
- b) the A- and D-Judges (for harmony, choreography, degree of difficulty, and music) and divide them by ??? to get the final % = **1 mark**

At the end the spectators have 2 understandable marks.