

Referee Evaluation Test at JWC 2017



Used Data



- Evaluators:
 - 3 only Fighting, rest did not return sheets¹
 - 4 Coaches → 2 only Duo, 2 Fighting
 - 6 Athletes → 1 only Duo, 2
- Evaluations (24 Referees):
 - 7 Evaluations Duo
 - 73 Evaluations Fighting
 - 2 Evaluations Ne Waza
 - Not evaluated

¹Many referee had no free time due to missing referees

Results Duo



- No evaluation sheet existed
 - Creation of an evaluation sheet
- Only Coaches & Athletes returned Duo Sheets
- Uncertainties:
 - Statistical: 35%
 - Systematic:
 - Different perspectives: 10%*
(0.5 Points)

**Guessed value*



Evaluation Sheet



*New Version, including explanations

Referee ID:		Evaluator: <input type="checkbox"/> Referee <input type="checkbox"/> Coach <input type="checkbox"/> Athlete			Country Evaluator:						
Category ¹ :		Red ² :		Blue ² :		Result ³ :		Agree ⁴ : <input type="checkbox"/> Yes <input type="checkbox"/> No			
	Points Referee ⁵	Own Points ⁶		Powerful Attack ⁷	Reality ⁷	Control ⁷	Effective-ness ⁷	Attitude ⁷	Speed ⁷	Variety ⁷	Comments
Series A (red)			+								
			0								
			-								
Series A (blue)			+								
			0								
			-								
Series B (blue)			+								
			0								
			-								
Series B (red)			+								
			0								
			-								
Series C (red)			+								
			0								
			-								
Series C (blue)			+								
			0								
			-								
Series D (blue)			+								
			0								
			-								
Series D (red)			+								
			0								
			-								

Comments:

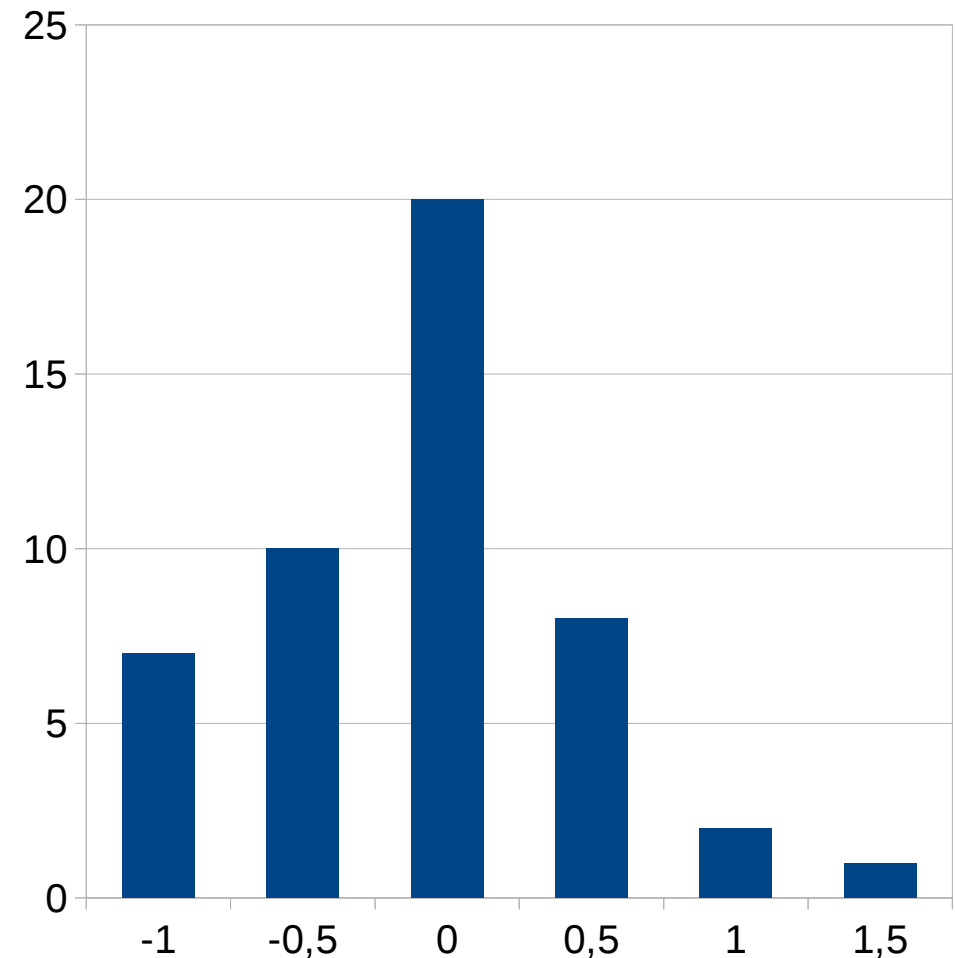


¹Men,Women,Mixed + Age (U18,U21,Sen) ²Country from Red/Blue ³Results of Match in Points ⁴Do you agree with the winner? If no please comment why ⁵Write Points the Referee gave ⁶Write points you would have given ⁷Give +,0,- for each criteria ©JJIF Technical Comission &Claudia Behnke

Difference in given points



- Observable:
 $P(\text{Eval}) - P(\text{Ref})$
 - 0: Referee and Evaluator gave same points
- $\sigma = -0.1$ $\text{Var} = 0.3$
 - On average the coaches would give little less points
 - Variance of points is within systematic uncertainties



Scoring in the Duo System



Section 23 Criteria of Judging ⁽³⁰⁾

- a. The jury shall look for and judge the following:
 1. Powerful attack
 2. Reality
 3. Control
 4. Effectiveness
 5. Attitude
 6. Speed
 7. Variety
- b. The overall score should give more importance to the attack, and to the first part of the defence.
- c. Atemis must be powerful, with good control and given in a natural way considering possible follow up.
- d. Throws and takedowns shall include breaking the opponents balance and be efficient.
- e. Locks and strangulations must be shown to the jury in a very obvious and correct way, with tapping by Uke.
- f. Both the attack and the defence shall be executed in a technical and realistic way.

Scoring in the Duo System



- Evaluators gave +,0,- on the 7 criteria
 - Expectation: a combination of +, - and 0 leads to a total amount of points
 - Put 50% more on powerful attack (See rules)

Powerfull Attack	Reality	Controll	Effctive ness	Attitude	Speed	Vairety	Points
+	+	+	+	+	+	+	10
-	-	-	-	-	-	-	0
0	0	0	0	0	0	0	5

Evaluation Sheet*



*New Version, including explanations

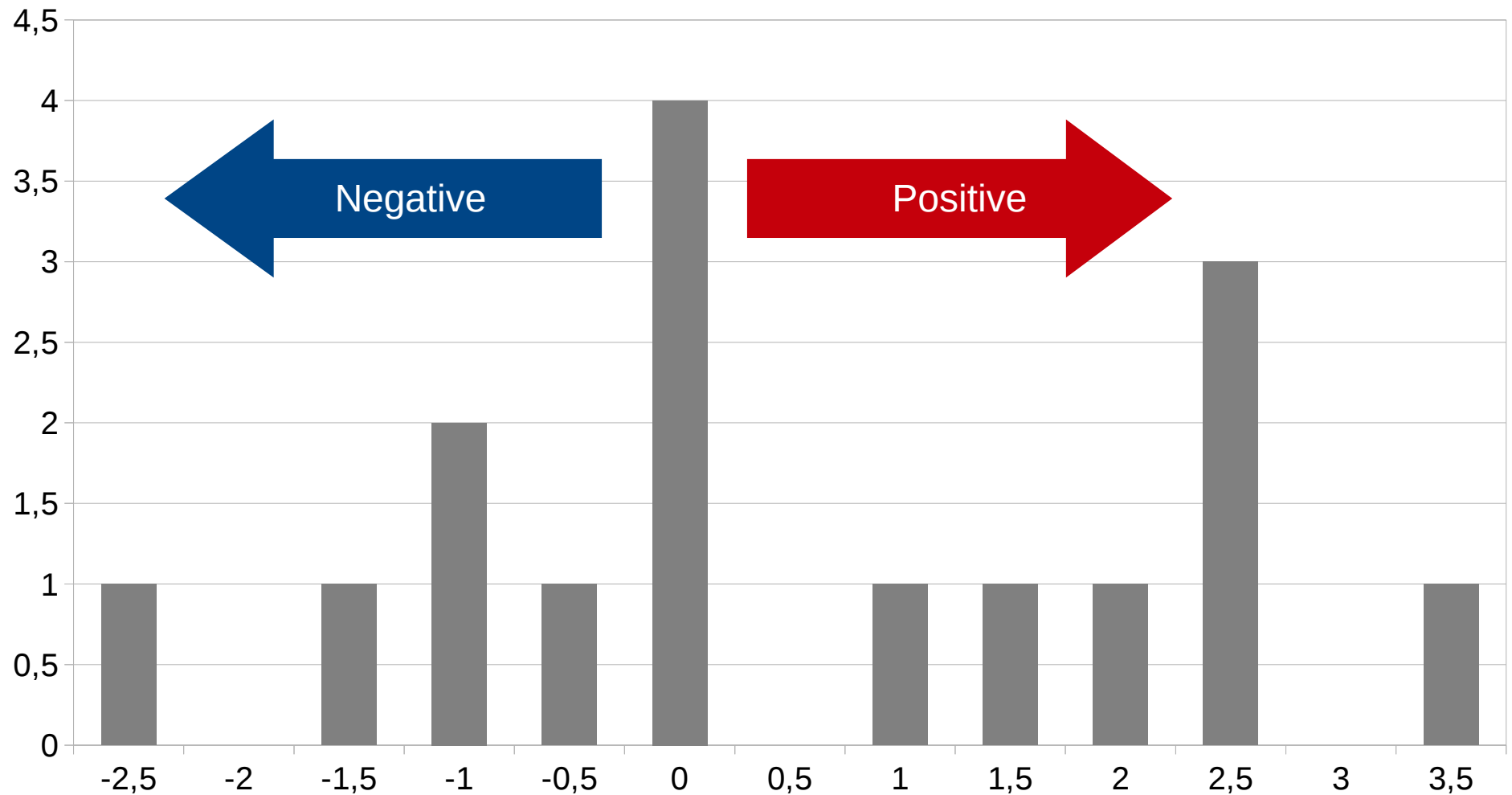
Referee ID:		Evaluator: <input type="checkbox"/> Referee <input type="checkbox"/> Coach <input type="checkbox"/> Athlete			Country Evaluator:						
Category ¹ :		Red ² :		Blue ² :		Result ³ :		:		Agree ⁴ : <input type="checkbox"/> Yes <input type="checkbox"/> No	
	Points Referee ⁵	Own Points ⁶		Powerful Attack ⁷	Reality ⁷	Control ⁷	Effective-ness ⁷	Attitude ⁷	Speed ⁷	Variety ⁷	Comments
Series A (red)			+								
			0								
			-								
Series A (blue)			+								
			0								
			-								
Series B (blue)			+								
			0								
			-								
Series B (red)			+								
			0								
			-								
Series C (red)			+								
			0								
			-								
Series C (blue)			+								
			0								
			-								
Series D (blue)			+								
			0								
			-								
Series D (red)			+								
			0								
			-								

Comments:

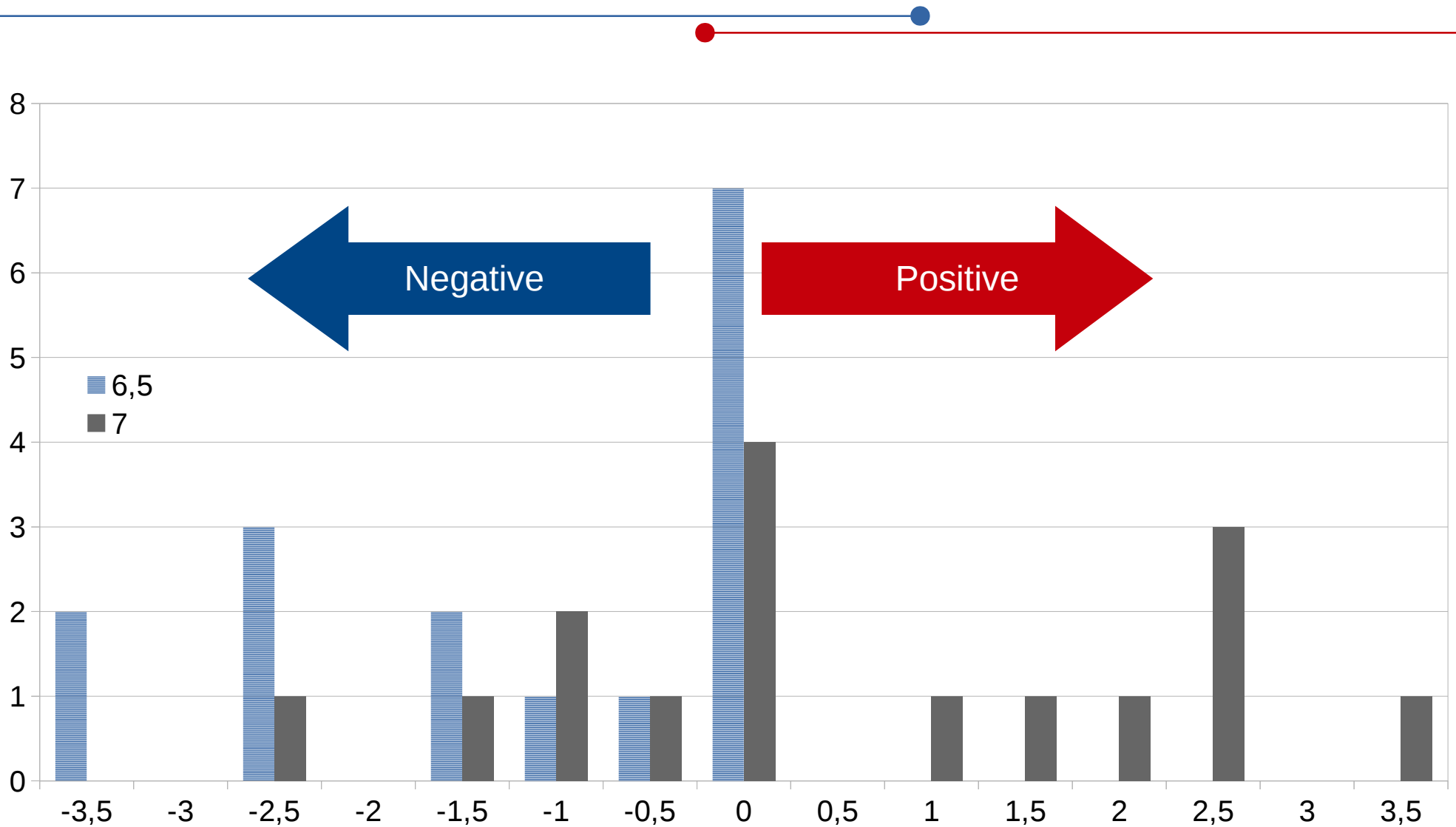


¹Men,Women,Mixed + Age (U18,U21,Sen) ²Country from Red/Blue ³Results of Match in Points ⁴Do you agree with the winner? If no please comment why
⁵Write Points the Referee gave ⁶Write points you would have given ⁷Give +,0,- for each criteria ©JJIF Technical Comission &Claudia Behnke

Example: 7 Points



Example: 6.5 & 7 Points



Conclusion Duo

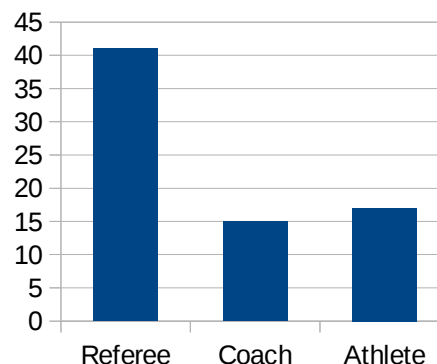


- Coaches & Athletes and Referees agree in given points within uncertainties
- No clear structure in point giving
 - Trends but not conclusive
- More data needed

Results Fighting



- Using the existing evaluation sheet
- 73 Returned sheets:
 - 41 from referees
 - 15 from coaches
 - 17 from athletes
- Uncertainties
 - Statistical: 12%
 - Systematical: *
 - 10% perspective
 - 5% biased due to athlete from own country

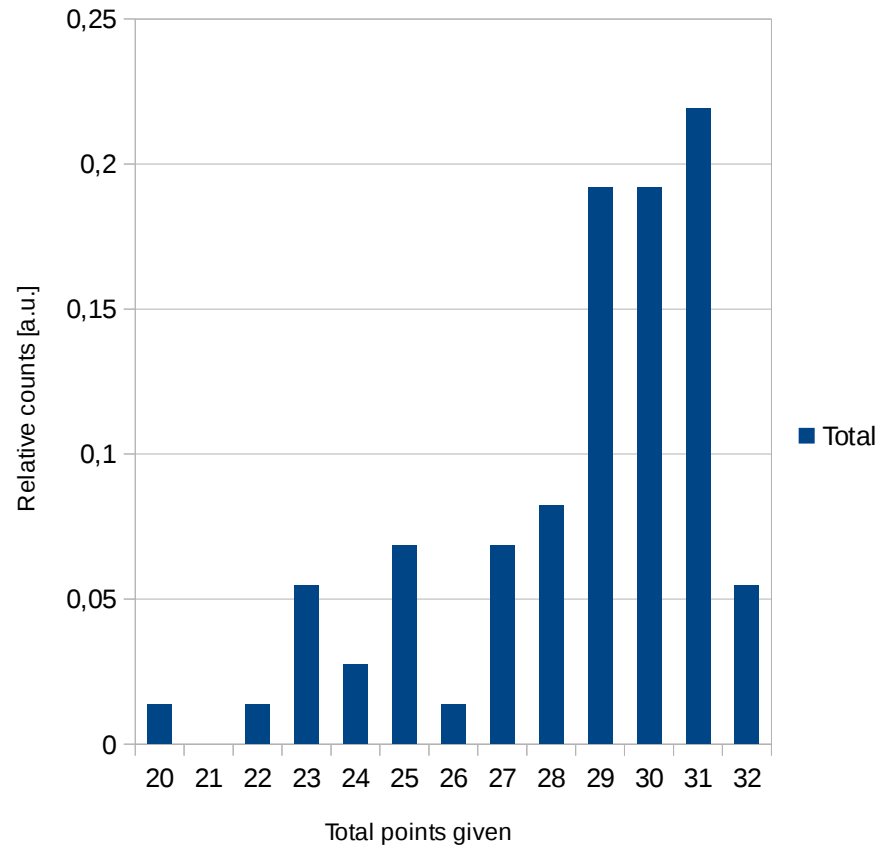


**Guessed value*

The average referee*



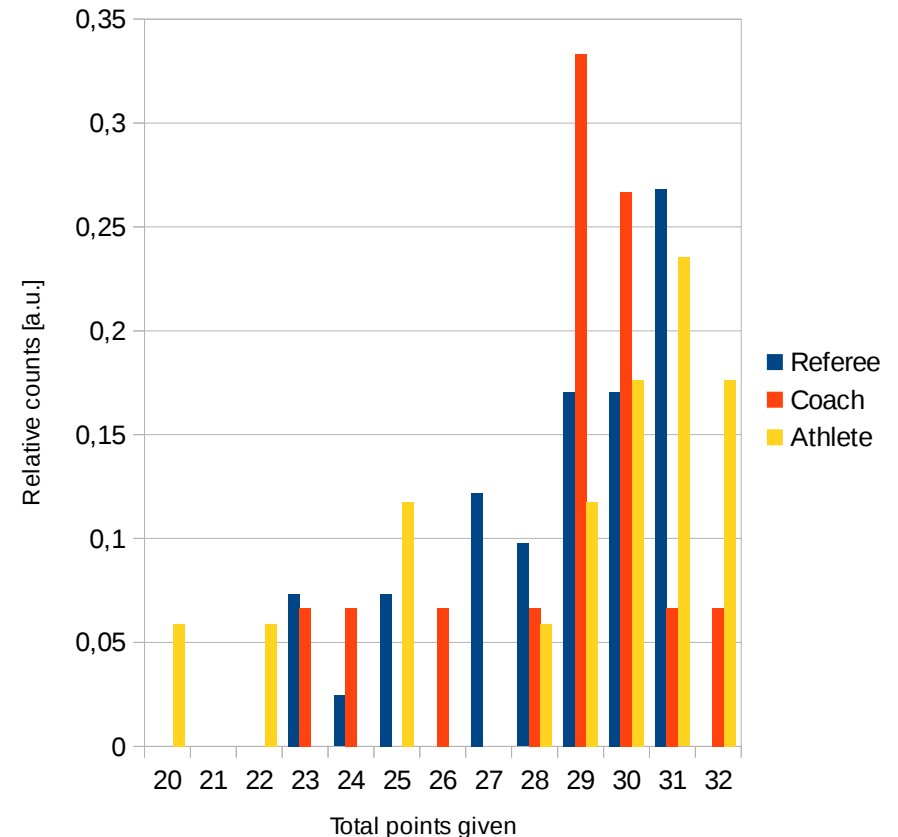
- *All referees together
- Average referee gets 28.5 ± 2.7 Points
- Does this value change for Different groups?



The average referee*



- *All referees together
- Average referee gets 28.5 ± 2.7 Points
- Does this value change for Different groups?
 - Referees: 28.5 ± 2.5
 - Athletes: 28.7 ± 3.6
 - Coaches: 28.6 ± 2.4
- All values agree
 - Athletes vary more



Evaluation sheet*

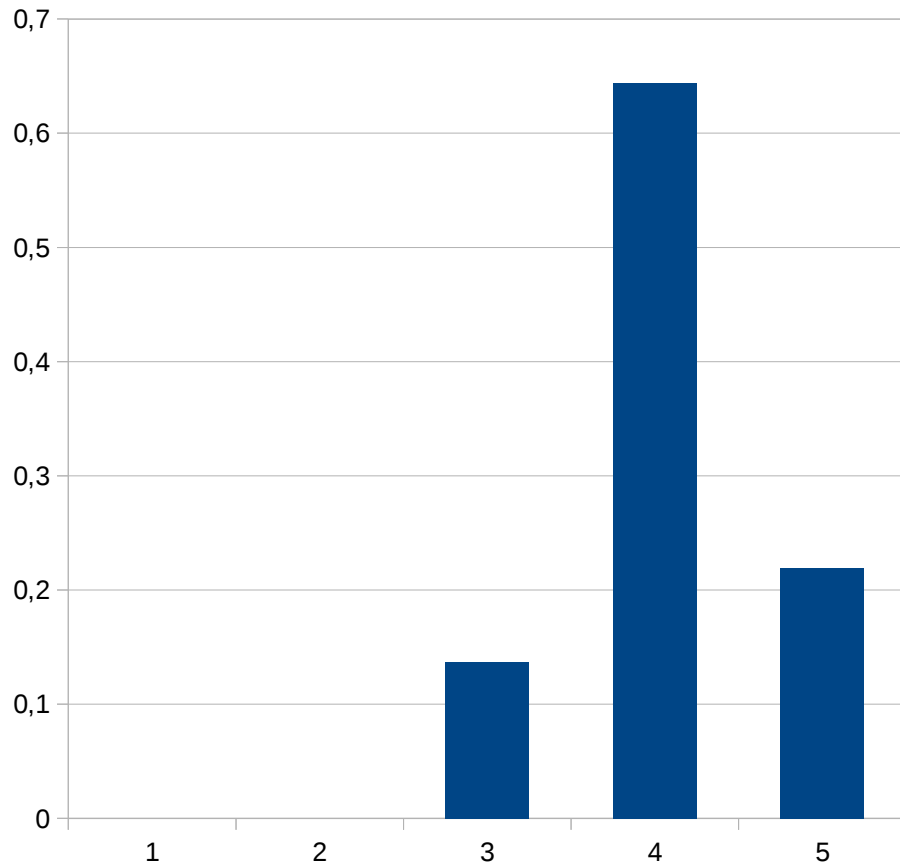


Score giving	Many faults, a lot of Points missed	1P					
		2P					
	Gives mostly the right score	3P					
		4P					
	Gives almost always the right score	5P					
Quickness in Score giving	To slow / fast	1P					
		2P					
	Sometimes to slow /fast	3P					
		4P					
	Good timing in giving the points	5P					
Control of the match	Lost the control of the match	1P					
		2P					
	Some uncontrolled actions of the fighters	3P					
		4P					
	Good Control of the match	5P					
Application of the rules	A lot of faults	1P					
		2P					
	Some faults	3P					
		4P					
	Good application of the rules, no faults	5P					
Flow of the match	A lot of unnecessary interruptions	1P					
	Some interruptions	2P					
	Only interruptions, if it's necessary	3P					
Contact with SR	Bad contact to the SR	1P					
	Sometimes	2P					
	Always in good contact to the SR	3P					
Movement and Bodylanguage	Often an unfavourable position as referee	1P					
	Sometimes an unfavourable position	2P					
	Always in a good, favourable position	3P					
Signs & voice	Misses a lot if signs, unclear signs / voice	1P					
	Misses sometimes the signs, unclear voice	2P					
	Uses always right & clear signs / voice.	3P					

- Used for referee evaluation since years
- 8 criteria that can give up to 32 points
- One referee is evaluated several times
- Results of each criteria will be compared now

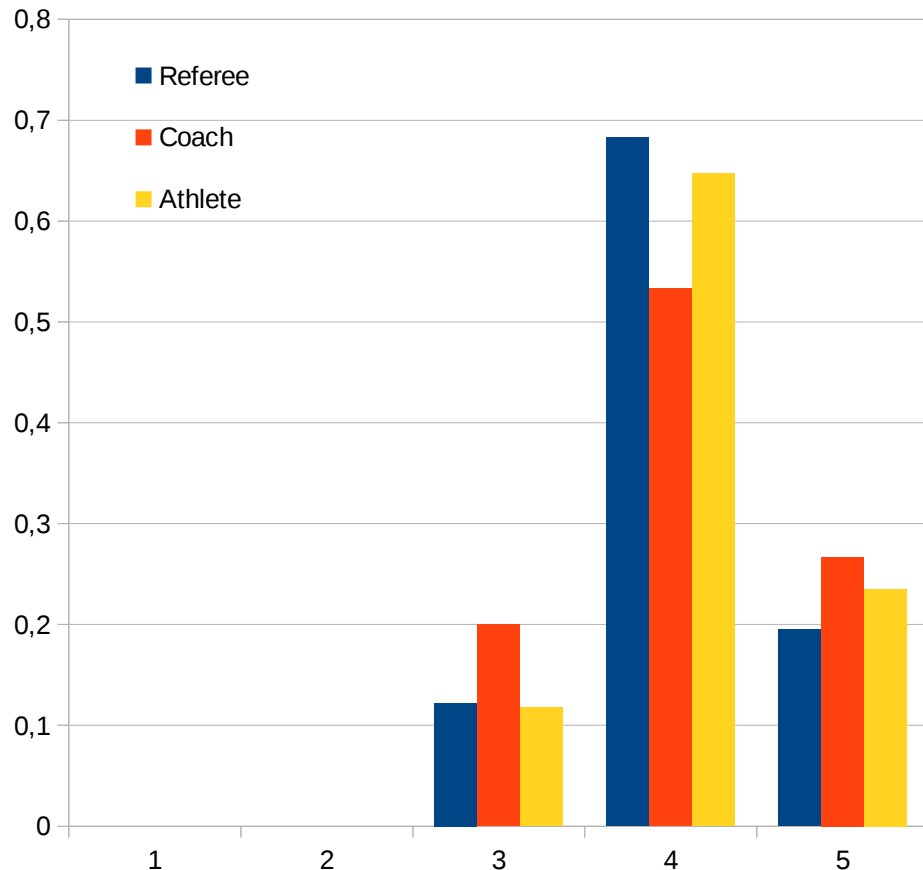
*Developed by referee commission

Average per criteria



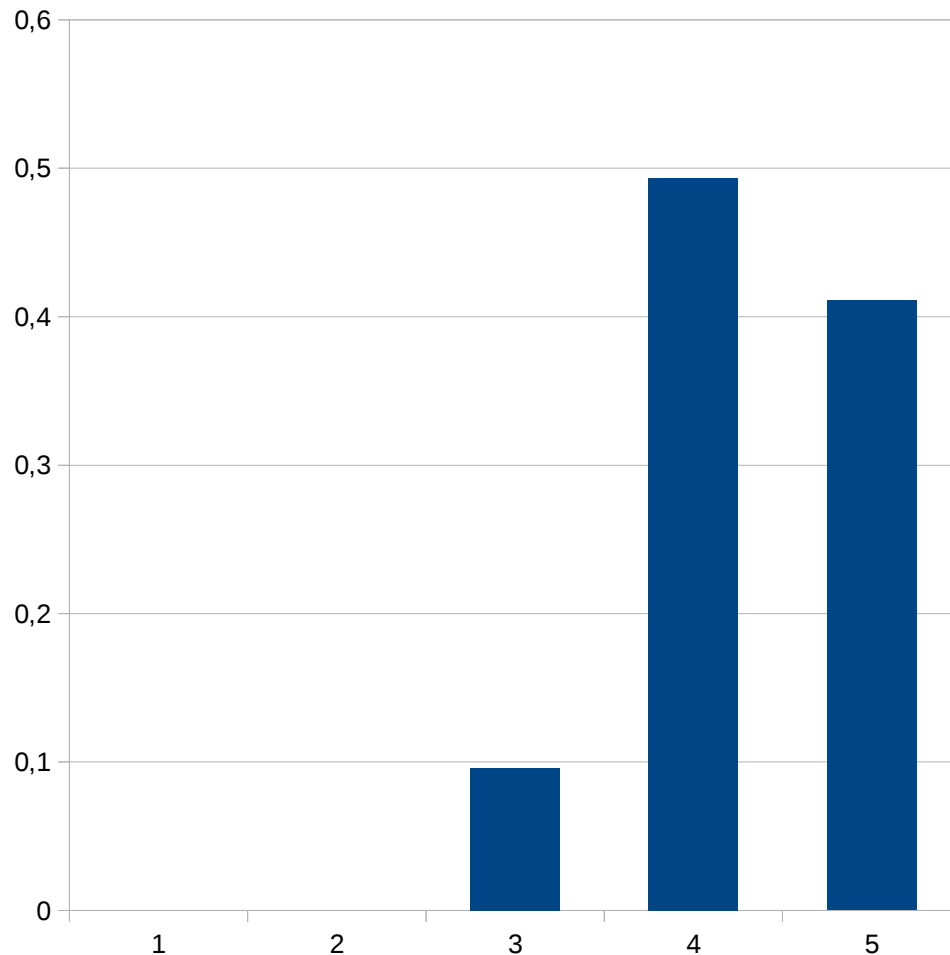
- Score giving criteria:
 - 1P: Many faults, a lot of Points missed
 - 3P: Gives mostly the right score
 - 5P: Gives almost always the right score
- Only 15% sometimes misses points, rest very good!

Average per criteria



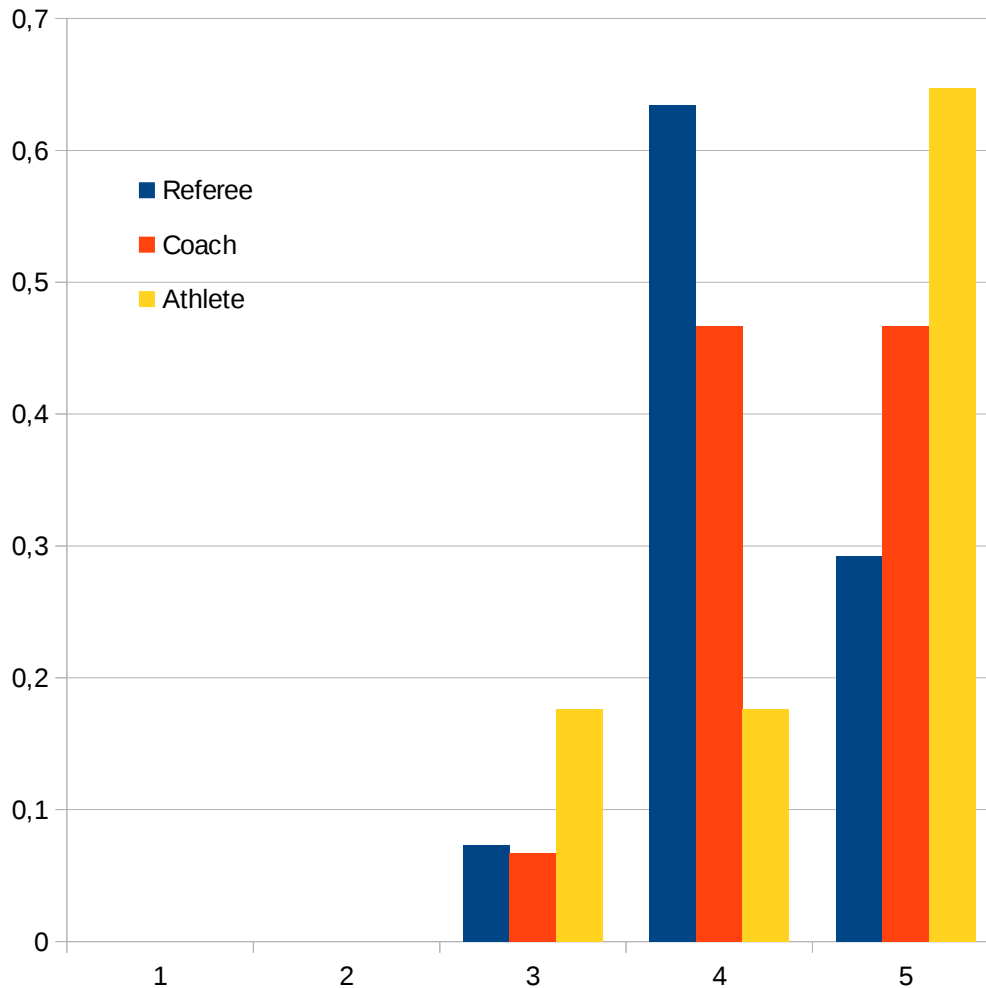
- Score giving criteria:
 - 1P: Many faults, a lot of Points missed
 - 3P: Gives mostly the right score
 - 5P: Gives almost always the right score
- Only 15% sometimes misses points, rest very good!
- Coaches see more missed points than athletes & referees

Average per criteria



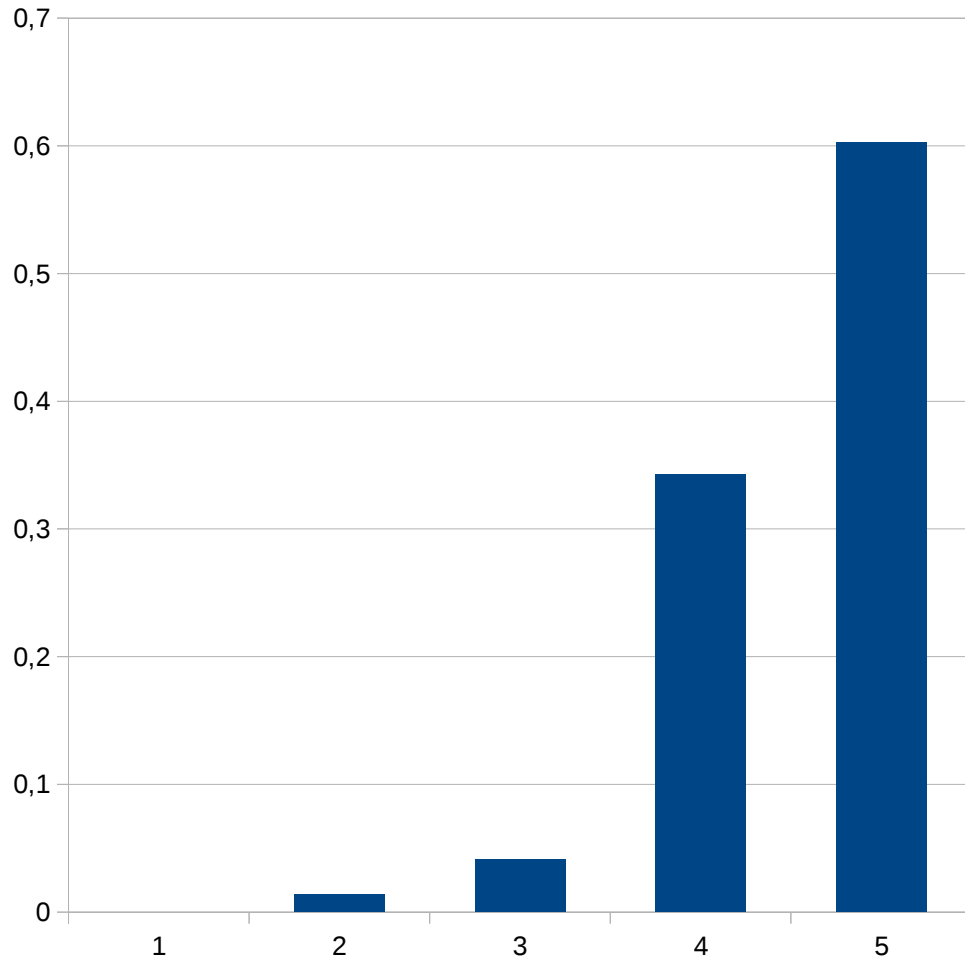
- Quickness in score giving:
 - 1P: To slow
 - 3P: Sometimes a little bit slowly
 - 5P: Very quick in giving the points

Average per criteria



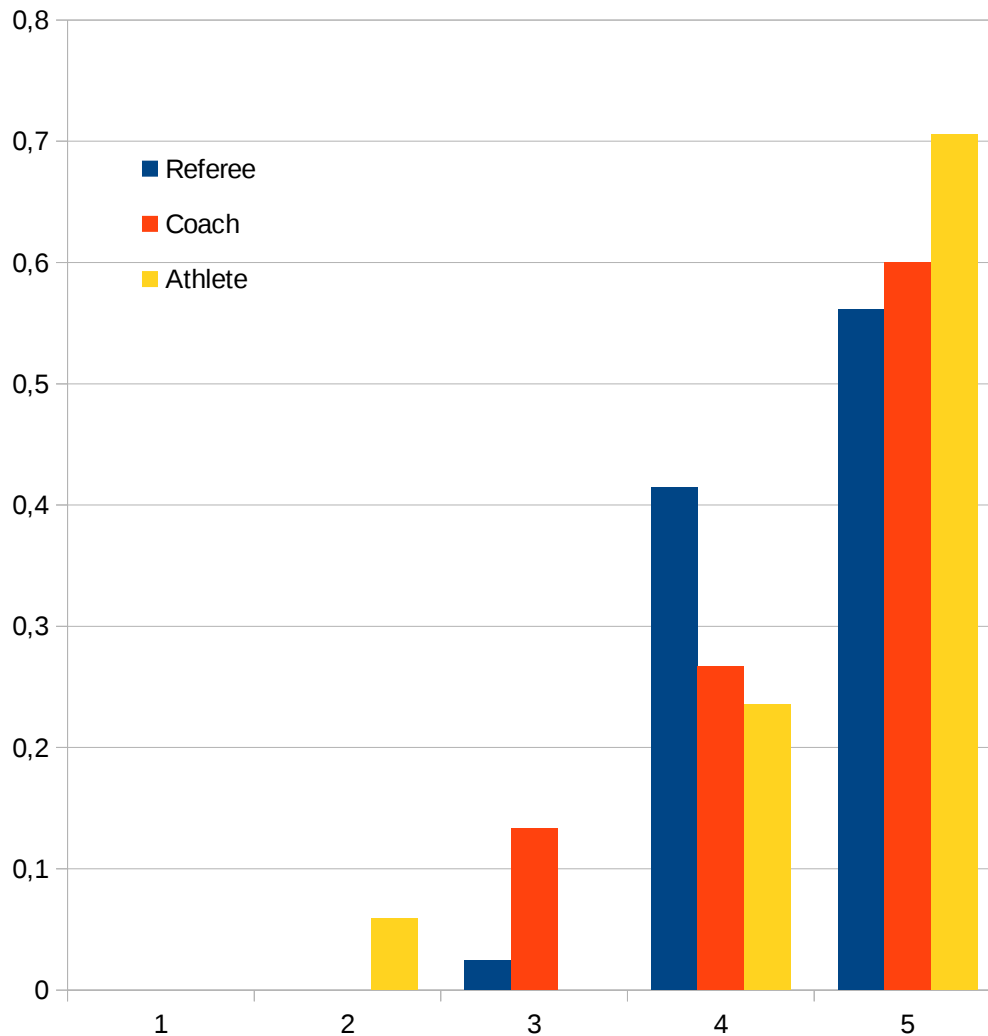
- Quickness in score giving:
 - 1P: To slow
 - 3P: Sometimes a little bit slowly
 - 5P: Very quick in giving the points
- Athletes and coaches give higher points than referees

Average per criteria



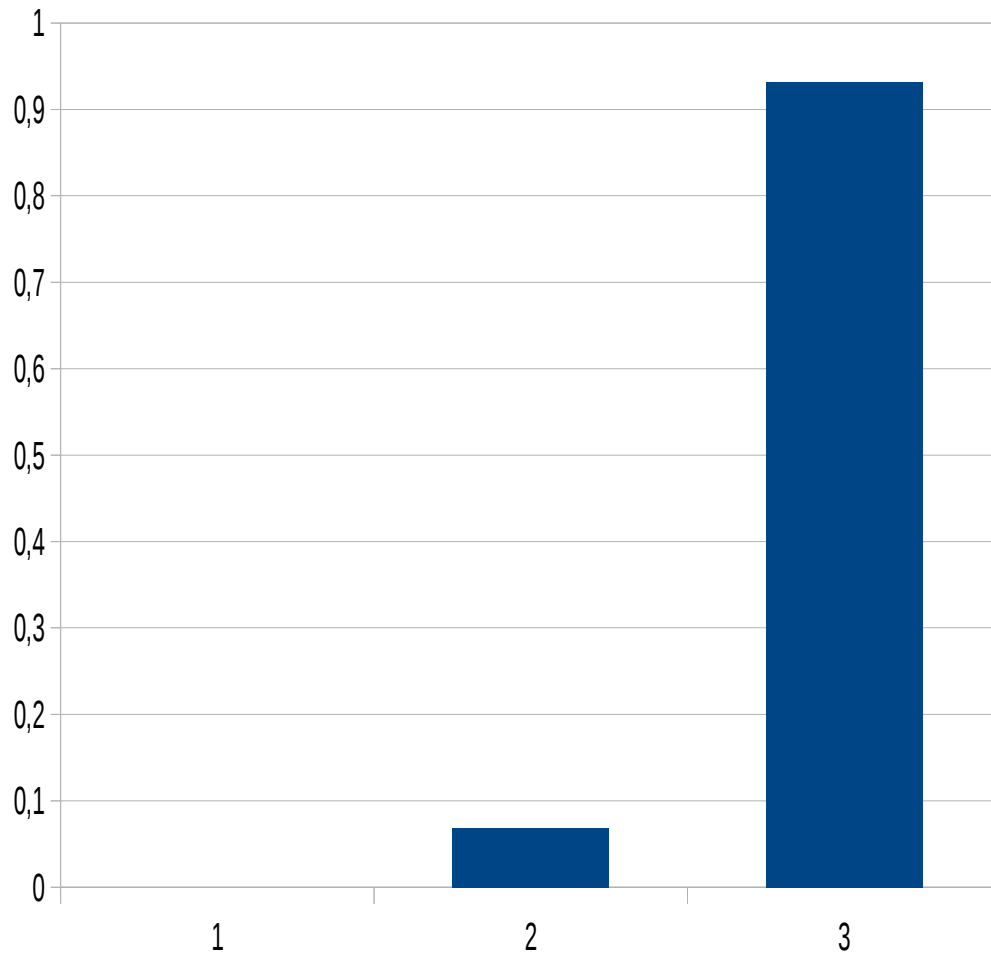
- Control of the match
 - 1P: Lost the control of the match
 - 3P: Some uncontrolled actions of the fighters
 - 5P: Good Control of the match

Average per criteria



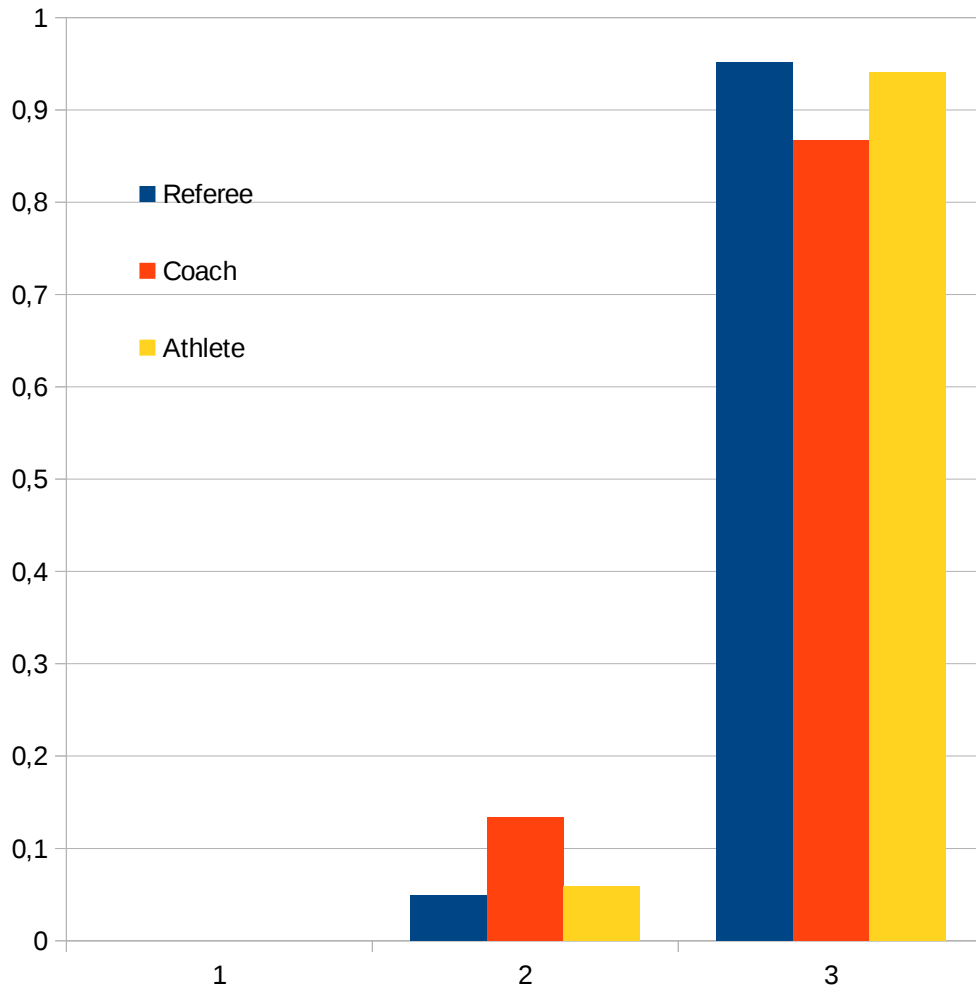
- Control of the match
 - 1P: Lost the control of the match
 - 3P: Some uncontrolled actions of the fighters
 - 5P: Good Control of the match
- Athletes and coaches give higher points than referees

Average per criteria



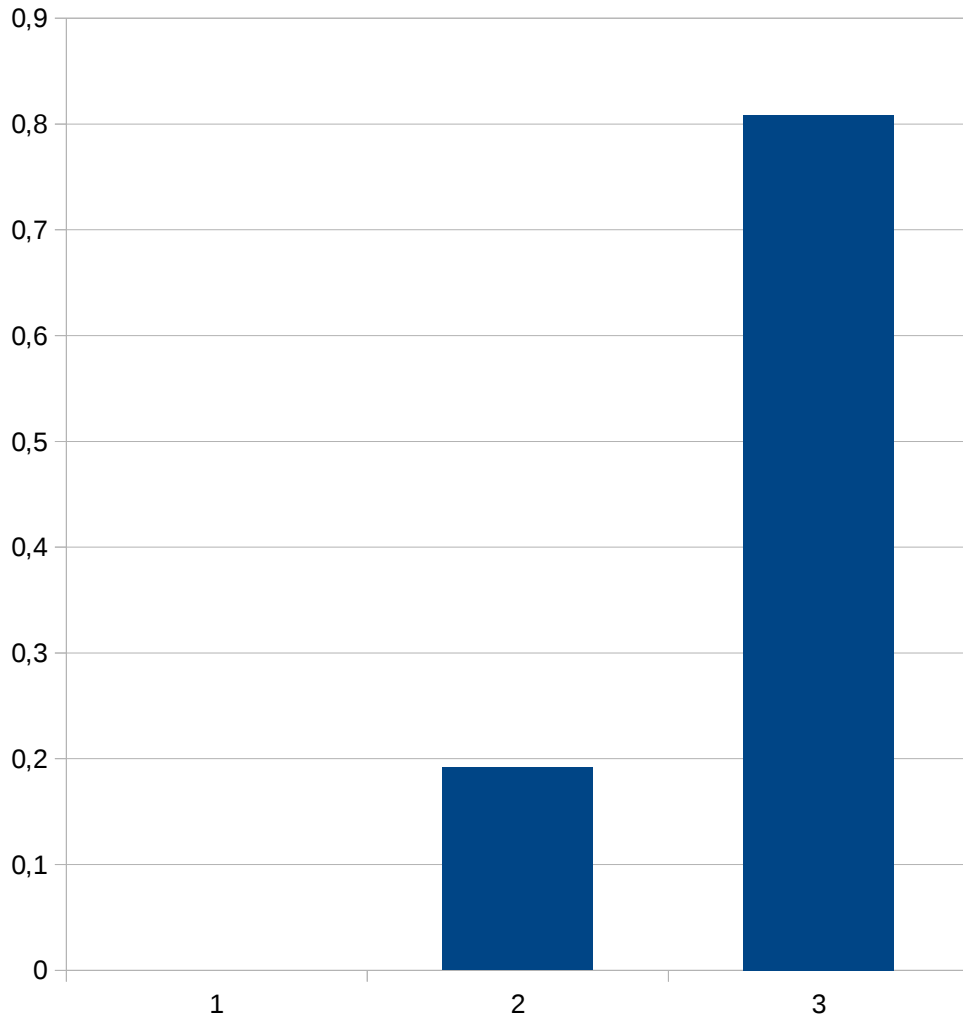
- Flow of the match
 - 1P: A lot of unnecessary interruptions
 - 2P: Some interruptions, which are not necessary
 - 3P: Only interruptions, if it's necessary

Average per criteria



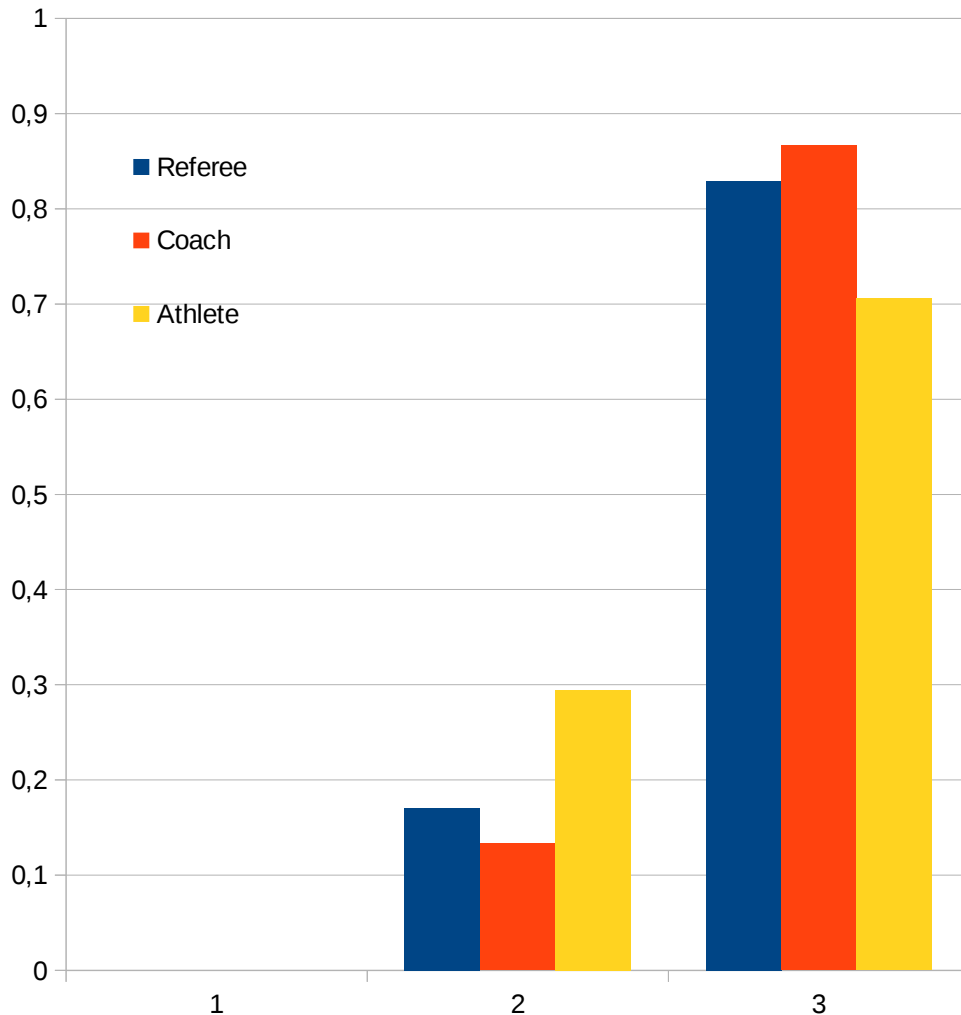
- Flow of the match
 - 1P: A lot of unnecessary interruptions
 - 2P: Some interruptions, which are not necessary
 - 3P: Only interruptions, if it's necessary
- Coaches score slightly worse, Athletes and referee agree

Average per criteria



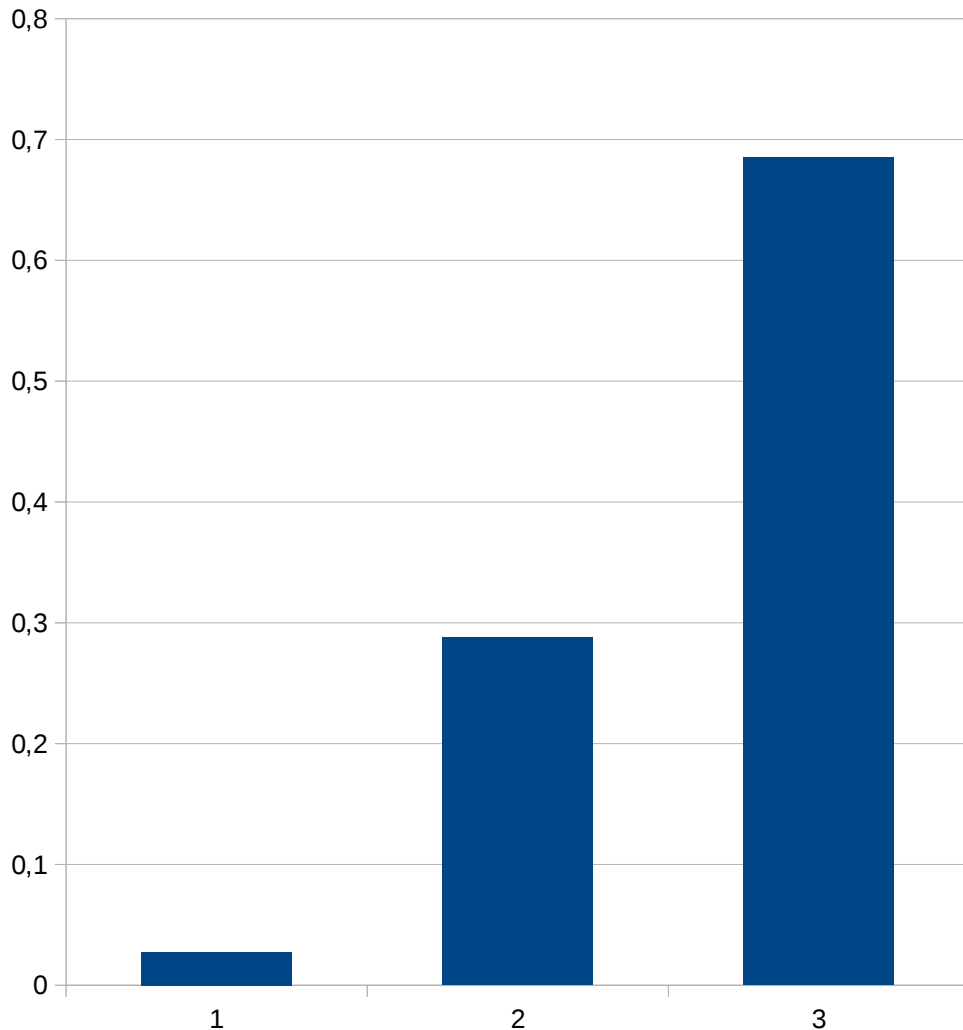
- Contact with SR
 - 1P Bad contact to the SR
 - 2P Sometimes bad
 - 3P Always in good contact to the SR

Average per criteria



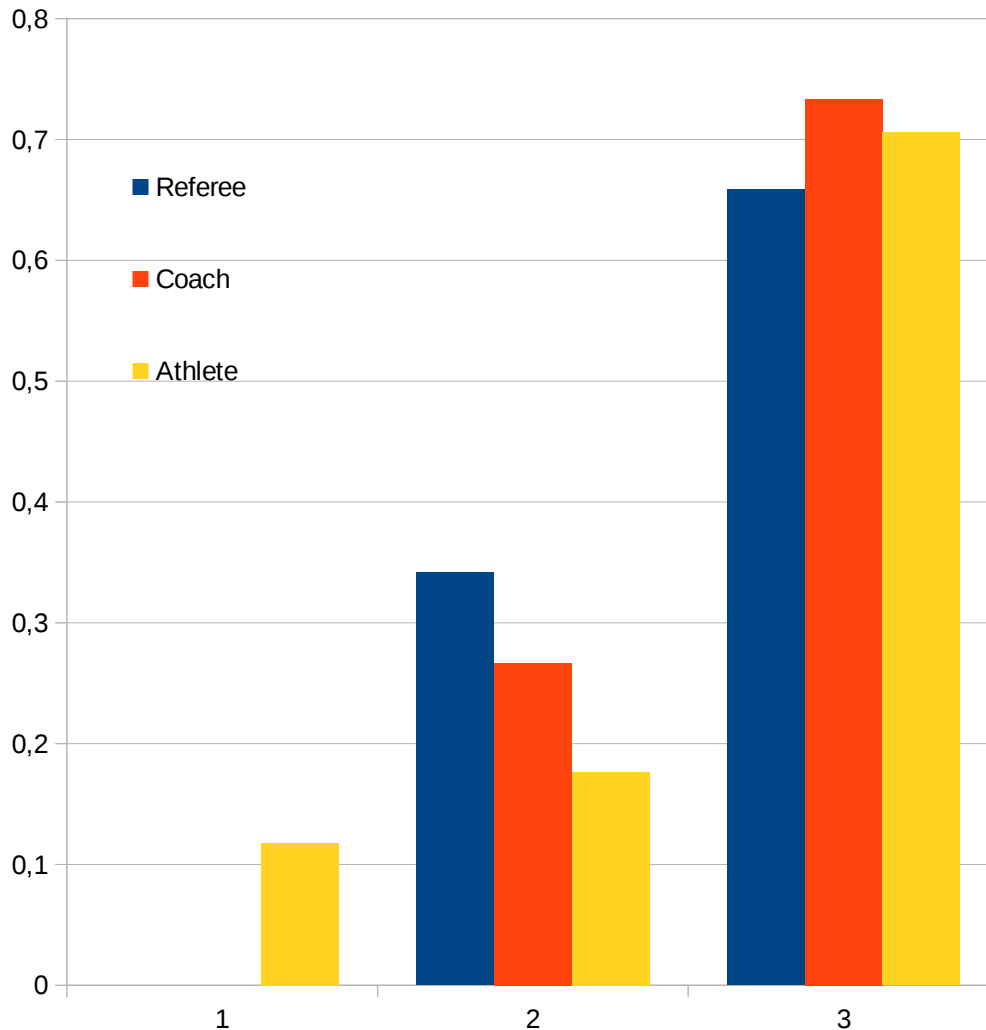
- Contact with SR
 - 1P Bad contact to the SR
 - 2P Sometimes bad
 - 3P Always in good contact to the SR
- Athletes give worse points
 - Is the criteria clear?
 - Is it visible for athletes?
- Coaches and Referees agree

Average per criteria



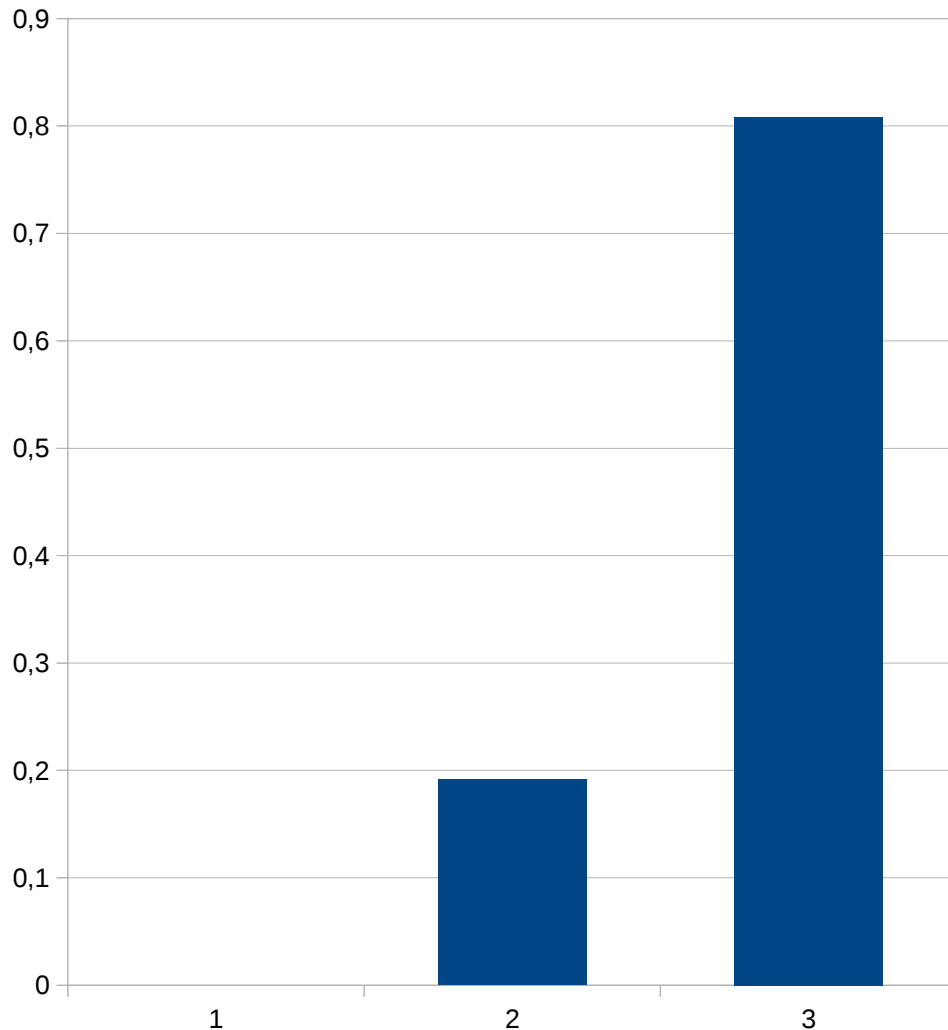
- Movement and Bodylanguage
 - 1P Often an unfavourable position as referee
 - 2P Sometimes an unfavourable position
 - 3P Always in a good, favourable position

Average per criteria



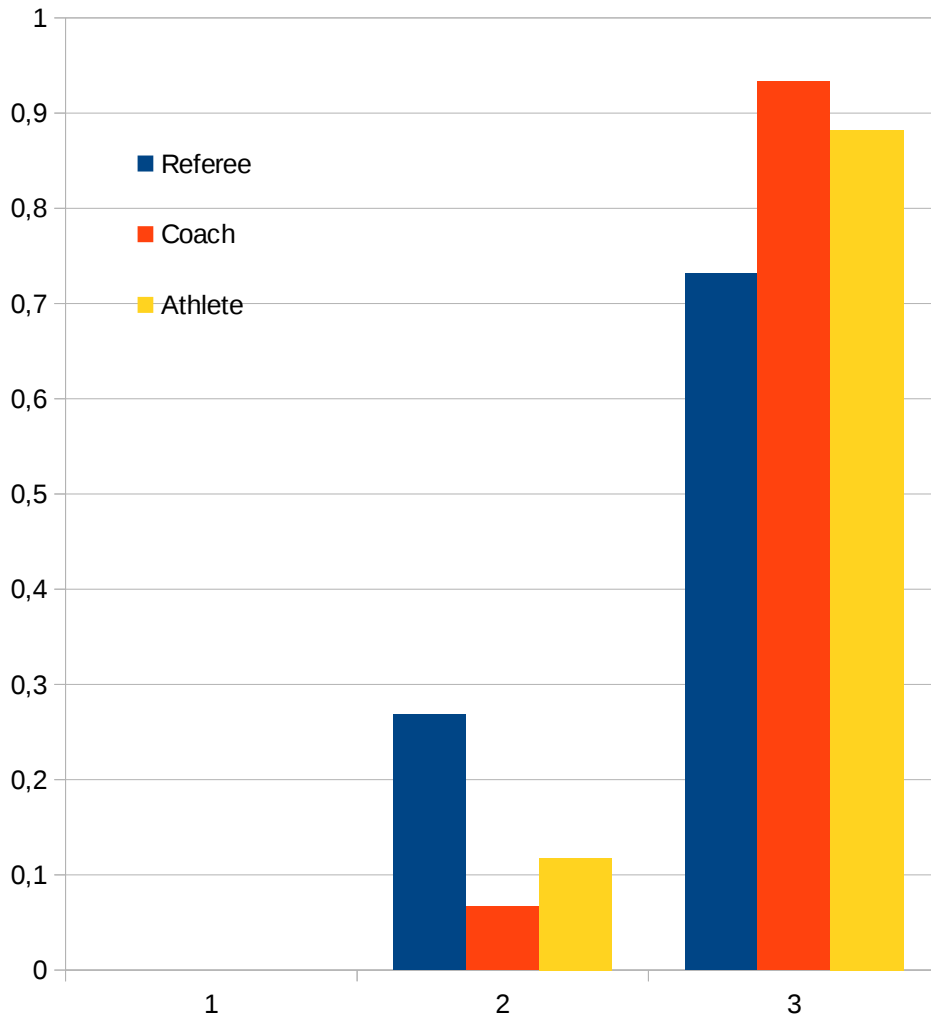
- Movement and Bodylanguage
 - 1P Often an unfavourable position as referee
 - 2P Sometimes an unfavourable position
 - 3P Always in a good, favourable position
- Athletes and Coaches give higher points than referees
- One referee was evaluated by one Athlete bad...
 - Significance?

Average per criteria



- Signs & voice
 - 1P Misses a lot if signs, unclear signs / voice
 - 2P Misses sometimes the signs, unclear voice
 - 3P Uses always right & clear signs / voice.

Average per criteria



- Signs & voice
 - 1P Misses a lot if signs, unclear signs / voice
 - 2P Misses sometimes the signs, unclear voice
 - 3P Uses always right & clear signs / voice.
- Coaches and Athletes give more points than referees

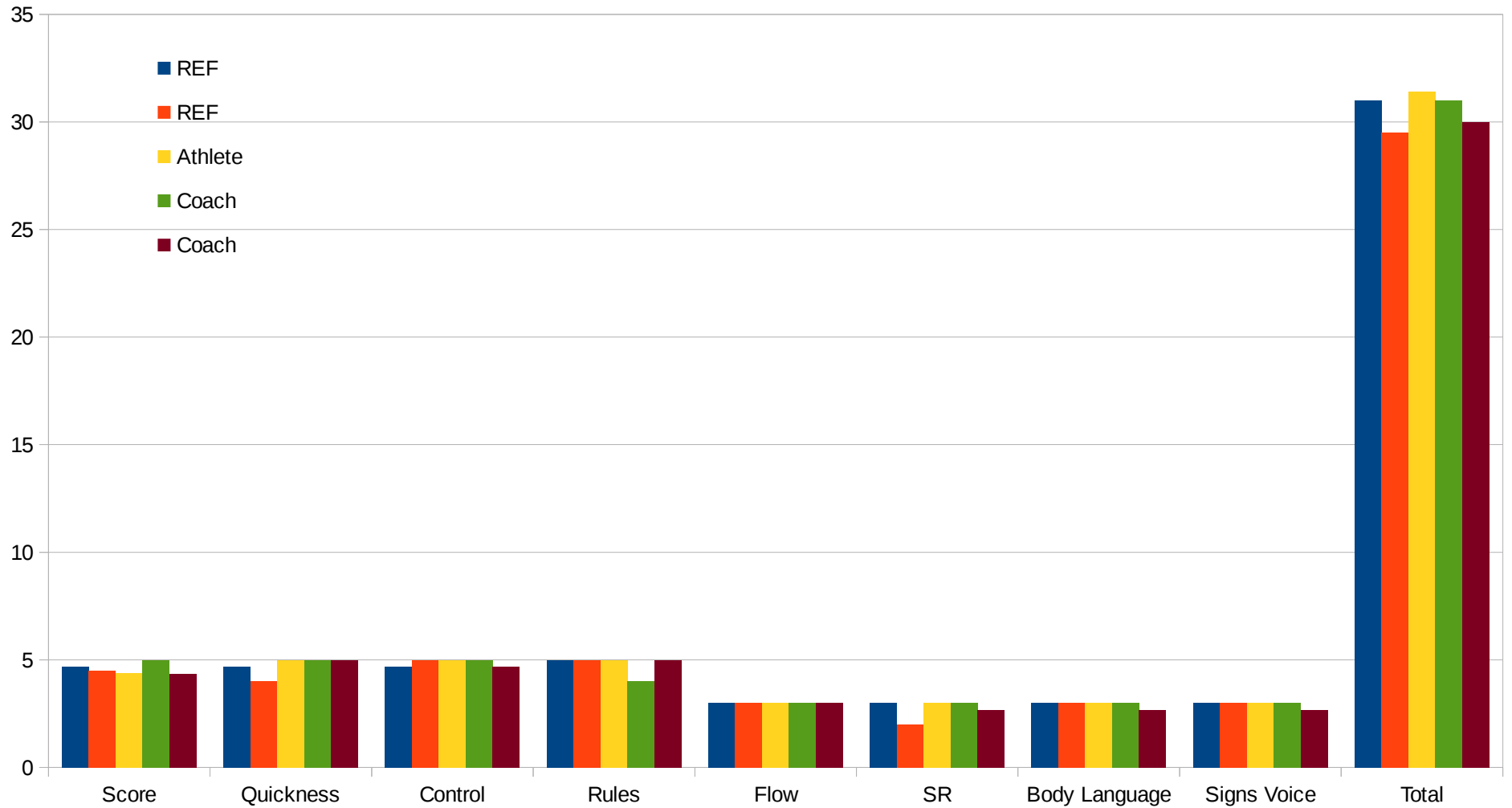
One single referee



- One referees was evaluated by 2 referees, 2 coaches and 1 athlete
 - Total 14 evaluations
- Uncertainties:
 - Statistical: 25%
 - Systematic*: 15% like before

**Guessed value*

One single referee



Conclusions



- Duo
 - There is no difference between the points given by the referee and the points given by an evaluator
 - The reason for point giving is not clear
- Fighting
 - Athletes, Coaches and Referees agree in their points
 - For total score
 - As well as within the different criteria
 - Some criteria need to be better explained
- Total
 - More data needed to be conclusive

Questions?



- Thanks to
 - all evaluators
 - all participating referees

Vous avez des questions?

У вас есть вопросы?

Noch Fragen?

Har du frågor?

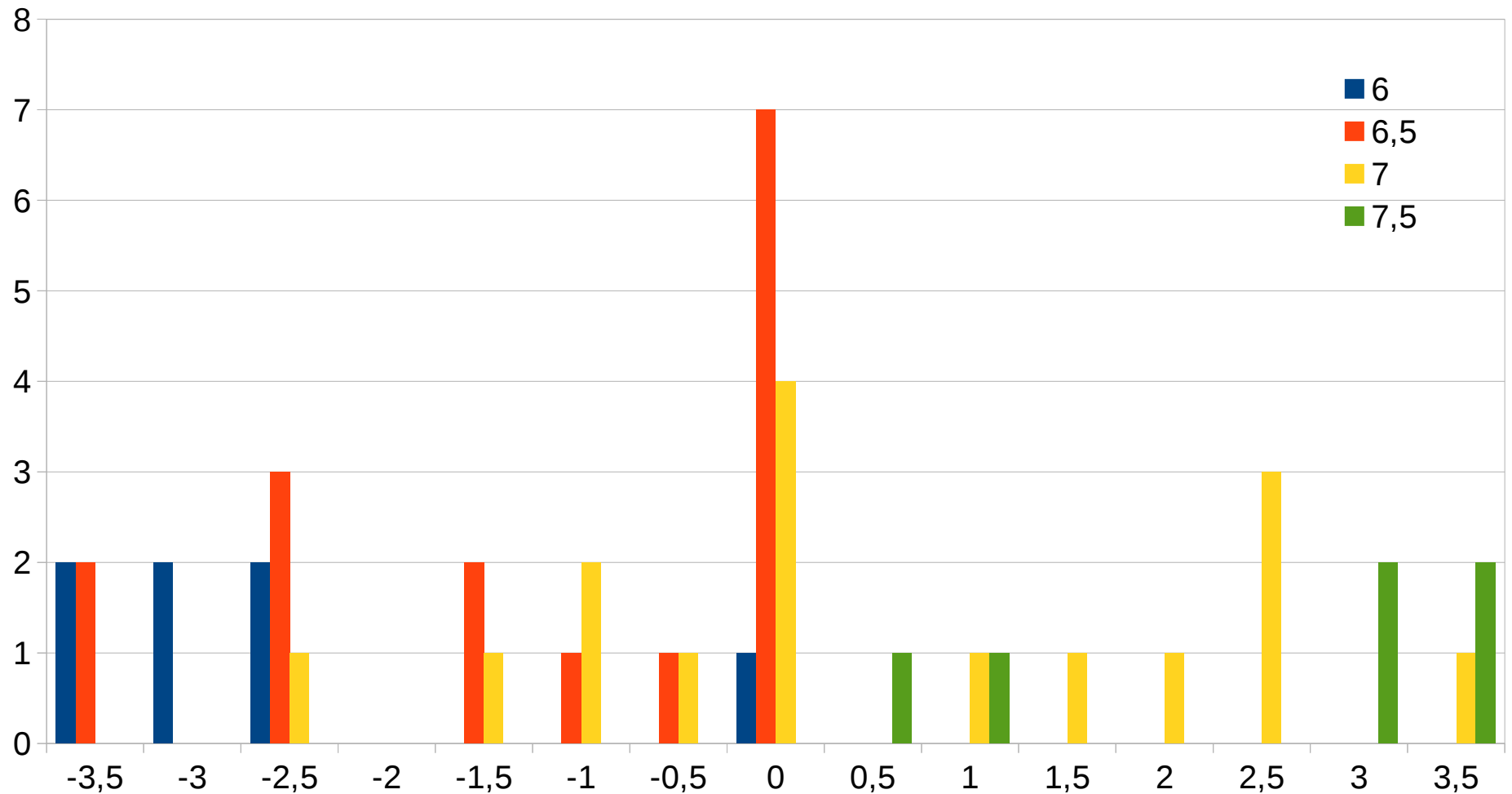
Heeft u nog vragen?

¿Tiene preguntas?

Bonus slides




Spread of Rank vs Points



Proposal for „new“ evaluation sheet



Referee ID:	Evaluator: <input type="checkbox"/> Referee <input type="checkbox"/> Coach <input type="checkbox"/> Athlete		Country Evaluator:				
 Fighting	Category ¹ :						
	Red ² :						
	Blue ² :						
	Result ² :						
	Agree ² :	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No	<input type="checkbox"/> Yes <input type="checkbox"/> No
Score giving	Many faults, a lot of Points missed	1P					
		2P					
	Gives mostly the right score	3P					
	Gives almost always the right score	4P					
Quickness in Score giving		5P					
	To slow / fast	1P					
		2P					
	Sometimes to slow /fast	3P					
Control of the match		4P					
	Good timing in giving the points	5P					
	Lost the control of the match	1P					
		2P					
Application of the rules	Some uncontrolled actions of the fighters	3P					
		4P					
	Good Control of the match	5P					
	A lot of faults	1P					
Flow of the match		2P					
	Some faults	3P					
		4P					
	Good application of the rules, no faults	5P					
Contact with SR	A lot of unnecessary interruptions	1P					
	Some interruptions	2P					
	Only interruptions, if it's necessary	3P					
Movement and Bodylanguage	Bad contact to the SR	1P					
	Sometimes	2P					
	Always in good contact to the SR	3P					
Signs & voice	Often an unfavourable position as referee	1P					
	Sometimes an unfavourable position	2P					
	Always in a good, favourable position	3P					
Signs & voice	Misses a lot if signs, unclear signs / voice	1P					
	Misses sometimes the signs, unclear voice	2P					
	Uses always right & clear signs / voice.	3P					

Comments: