

Referee Column: Australasian Taekwondo magazine.

By Kim Seng,
National Referee Chairman, Taekwondo Australia
WTF International Referee 1st class
Referee Chairman 2005 Oceania Championships
Referee Chairman 2005 Arafura Games



Evaluating referee performance for continuous improvement

“Referees are a necessary evil” says one of Australia’s team managers, when asked what he thought of referees. A player said that “Referees just do their jobs, they just penalize us and they do not give us enough points”. There are many more comments you hear all the time, some are excellent ones that complements the referees while others criticize the referees, fortunately there are more positives than negative comments.

The referees think some players and coaches out there make life hard for them when they officiate and do their jobs. Referees always try to do their best when they officiate. They remain impartial and they are in the unfortunate and an unenviable position where they judge and manage a fight, and they ultimately have to declare the result. The result is always unfavorable to one player in a match. The referees are always in the firing line. But such is the reward for being a referee. Referees must uphold the “law”, as they are the custodians of “policing” the competition rules and manage the fights in accordance to the competition rules of the day. They score the points as they confirm the legitimate attacks on legal scoring areas that have sufficient power and accuracy.

When you think about it a little more, the referees are there only to manage the fights and score the points that are the work of the players. When the player manages his techniques well and delivers his legitimate attacks in the right legal scoring areas and with the right amount of power and accuracy, the referee judge will score the point, there is to be no doubt. The referee does not score the point if the criteria for a point did not meet the criteria for a point score or if there is any doubt.

Similarly, if a player manages his fight well and does not commit any unacceptable match management tactics, then the referee has no reason to penalize the player. The referee has guidelines to manage the fight and if the player go outside of these guidelines, then of course the referee can penalize the player. Much depends on the severity or intentionality of the player’s actions, whether if the referee would either give a joo-eui verbal warning or a kyong-go penalty.

The WTF reforms on competition rules will make the sport more exciting with encouragement for faster paced tactical exchanges and better rewards (more points) for more difficult kicks that score and for players who can score quickly or abundantly. From 2006, Australia will implement the 7 points gap and the 12 points ceiling winning decisions.

Players and coaches will now have to work on new strategies to implement these new rules. The same goes for the sudden death rules, when implemented since April this year. After all it is the players who should be the ones who score those points and win the matches, and the referee is there to confirm those points for the players. It is not like the points are manufactured or denied by the referees. If it is there it will be scored.

Much as players and coaches embark on ways to ensure those points are scored, they must be sharp, strong, clear and precise. Referees will also undergo referee evaluation to ensure we are consistent in our match management and also consistent in our point scoring. In Australia we have begun to monitor scoring profiles and referee match management

evaluations. These are necessary so that we remain on track to maintain our standards and to keep improving ourselves. To remain in control of our referee management on the court, we now have referee evaluation conducted on all referees at a national event. These evaluations are used to monitor strengths and weaknesses of our referees and highlight areas for improvements where there are gaps to best practice. Where there are excellent skills, these will be used as benchmarks for other referees to follow. The criteria used for evaluation include referee conduct and attitude, referee hand signal techniques, application of competition rules, decision making, presentation skills and style, ring craft and match management and supervision or mentoring abilities. The table below shows the format for the evaluation. It can be used for any form of evaluation monitoring.

Referee Assessment Review

Assessor:

Event:

A+	A	B+	B	C+	C	D+	D	E+	E
10	9	8	7	6	5	4	3	2	1

Referee	Match Number	Referee performance profile (circle score) where appropriate										Categories being assessed	74.3	Comments (strengths & weaknesses)	Referee Signature
		10	9	8	7	6	5	4	3	2	1	Referee conduct & attitude	9		
		10	9	8	7	6	5	4	3	2	1	Referee hand signals	8		
		10	9	8	7	6	5	4	3	2	1	Application of competition rules	7		
		10	9	8	7	6	5	4	3	2	1	Decision making	8		
		10	9	8	7	6	5	4	3	2	1	Presentation skills & style	7		
		10	9	8	7	6	5	4	3	2	1	Ring craft and match management	6		
		10	9	8	7	6	5	4	3	2	1	Supervision and-or mentoring ability	7		
												Net total	7.4		

The main use of these evaluations is to be a transparent and feedback tool for the referees and enables referees to lift the standards to benchmark levels. Much the same as players and coaches try to lift their standards, so too will the referees. We cannot remain stagnant and we should always be looking at ways to get better all the time. Plus this gives us a method to quantifiably measure our performances and keep in control of our actions and skill levels.

Another implementation will be to measure each referee's scoring profile. I have devised a method of measuring referee judging and their scoring profiles using each judge's scoring incidences in a match. Each scoring incidence is measured against the final score that had been agreed by 2 judges (3 judges from 2006) where the points are validated and displayed. eg if judge #1 scores 3 times in a round and the score validation is 5 points, then mathematically judge #1 is 60% accurate for that round. If we did this for the 3 rounds, we may end up with 60% accuracy overall eg. One match is insufficient to determine if the judge #1 is actually a 60% accurate judge. A larger sample needed to be used to measure his consistency over a larger sample. Say for example, this judge #1 scores 60%, 75%, 80%, 75%, 80%, 75%, 80%, 75%, 80%, 80% over 10 matches, then we can statistically conclude that he is more like a 75%-80% accurate judge and the 60% in the first match is an exception.

Do not be alarmed that accuracy of match scoring is 80% as being not perfect. As we all know there is no perfection in scoring, due to many factors such as judges positioning being obscured and not able to confirm the point, such as the impact area is indeterminable, such as the time span between pressing the button is outside of the 1 second allowance and other factors. The useable data is to measure the referee judge's profile over time and check for consistencies and repeatability. To remain in control, the referee judge has to repeat his scoring profile consistently and any drop in profiling could be a sign of fatigue or any incidence for further review.

There are many uses of this scoring profiling such as

- to highlight any misunderstanding of point scoring criteria. The WTF and all referees are very keen to ensure there is unified criteria for scoring,
- to highlight any major under scoring or high scoring referee scoring profiles. Statistically it is better to have an over scoring profile than an under scoring profile. Any differences or major variances in scoring profile can highlight different criteria understanding of point scores and this can be picked up early and corrected,
- to highlight good on target scoring profiles, which demonstrates very good control,
- erratic scoring profiles, this demonstrates a scoring profile which is out of control and needed major attention and correction,
- uncontrolled scoring profiles, ie scoring profiles that vary from one extreme to another at different times. This demonstrates exceptional cases and requires further review.

The table below show examples of what these scoring profile look like.

Scoring test profiling				accuracy scoring profile %			consistency scoring profile %		
		grade	total mark	over score	on target	under score	over score	on target	under score
		tested		over score	on target	under score	over score	on target	under score
<i>do not fill-in any data in this</i>									
AA	ZZ	3	76.9%	36%	43%	21%	29%	57%	14%
BB	YY	2	82.1%	36%	57%	7%	36%	57%	7%
CC	XX	4	74.4%	14%	43%	43%	14%	43%	43%
DD	WW	2	84.6%	21%	64%	14%	21%	64%	14%
EE	VV	3	71.8%	21%	29%	50%	14%	36%	50%
FF	UU	2	79.5%	14%	43%	43%	14%	43%	43%
GG	TT	1	69.2%	43%	36%	21%	43%	36%	21%
HH	SS	2	84.6%	21%	57%	21%	21%	57%	21%
II	RR	1	87.2%	21%	57%	21%	14%	71%	14%
JJ	QQ	2	82.1%	14%	57%	29%	14%	57%	29%

The use of scoring profile and referee evaluation are powerful tools that can be used to reduce variability and maintain or improve controllability. The end result will be continuous improvement as variabilities are narrowed progressively until we reach scoring profiles and referee performance profiles that are consistently in control.

Editor's note: readers are more than welcome to ask any questions about refereeing, competition rules or referee management.