

Analysis of Reliability and Validity in Refereeing of Kumite in Karate-Do

**Dissertation presented towards Master of Sports Science Degree
with specialisation in Sports Training**

Fernando José de Carvalho Fidalgo

Master's Supervisor:

Félix Luís de Lima e Cunha Hopffer Romero PhD

Acknowledgements

I would like to start by thanking my family, especially my parents for the way they always transmitted to me values of work, honesty and human comprehension. To my sons Pedro e João for understanding their father's "absence".

To my wife, Isabel Pata Fidalgo, for the determining support writing this paper, for the comprehension, affection and specially the help in its final revision.

To Professor Doutor Félix Romero, supervisor of this work, for the dedication, for the understanding, for the incentive, for the time made available, for the counselling and the transmission of knowledge.

To Shihan Mário Águas, in his quality of President and Technical Director of the Portuguese Shotokan Karate Federation (Federação Portuguesa de Karate Shotokan) and personal representative in Portugal of Kancho, Hirokazu Kanazawa, for the way he took the idea, supported, gave incentive to the realization of this study, and above all for his teaching over the last 25 years.

To Kancho Nobuaki Kanazawa, for his availability to be part of the experts commission.

To Nagata Sensei Luís Filipe and José Mário Mil-Homens, for their availability and support, both in the experts commission as well as in the production and analysis of the videos.

To Nagata Sensei Joaquim Reis and Maximino Dias for their availability and logistic support.

To the referees and competitors of the Portuguese Shotokan Karate Federation who were available to participate in this study.

To Professor Doutor Luís Carvalhinho for the incentive to apply for the Master's.

To Professor Doutor Carlos Silva for the support in the psychology area.

To Dr^a Michelle Sabino for her help in translating the abstract.

To the comrades at the Direction of the Portuguese Red Cross Delegation in Caldas da Rainha, Major João Machado and Captain Manuel Filipe, for replacing me in some functions and especially to Petty Officer João Mouro, for replacing me in the "Clic Senior" IT classes.

To the teachers at the Karate Shotokan Club in Caldas da Rainha, Rui Filipe e António Folgado for replacing me in many training sessions over the first year of the master's.

To all the teachers at ESDRM, who in the first year of this second cycle transmitted us fundamental knowledge, for us to follow our investigation projects in a more autonomous way, and also for their permanent availability.

To the labs at ESDRM for the availability of material for the data recordings.

To Ruben Anastácio for the translation of this work to english and Jane Ribeiro for its review.

To every person or entity who in some way participated and supported me in this project.

Thank you.

Resumo

A investigação centrada nos árbitros tem sido pouca ou escassa, sendo praticamente inexistente na arbitragem de kumite em Karate.

O presente estudo teve como objetivo, caracterizar o desempenho dos árbitros de Karate, na competição de kumite, quanto à sua validade, e ainda à fidelidade intra-observador e à fidelidade inter-observador.

A presente investigação assentou na observação, por parte de árbitros de Karate da FPKS, de um combate de Karate com a duração de 2 minutos. Desse combate foi realizado um vídeo com a duração de 8 minutos que reproduzia o combate observado quer da posição do árbitro, quer da posição de cada um dos juízes.

A recolha de dados foi efetuada em dois momentos, o primeiro com vista a analisar a validade e a fidelidade inter-observador e o segundo momento, com um intervalo mínimo de 60 dias para verificarmos a fidelidade intra-observador.

Para a análise da validade foi constituída uma comissão de peritos, com o propósito de se obter uma medida critério.

A observação do vídeo por parte da amostra foi feita individualmente pelos árbitros, em ambiente fechado, através do software Match Vision Studio, em que se fez o registo da observação.

Utilizámos o índice de Bellack, para aferir a fidelidade intra e inter-observador e a validade. Aceitando-se um resultado igual ou superior a 80% como garante da fidelidade (Aranha, 2007, Siedentop, 1983). Foi utilizado o Qui-Quadrado para verificar se havia diferenças significativas na validade entre árbitros nacionais e regionais e em função da posição de observação (árbitro ou juiz).

Os valores da fidelidade intra-observador (78,7%) e inter-observador (77,2%) estão próximos, o que nos diz que a variabilidade encontrada entre os diversos árbitros não é superior às variações que ocorrem num mesmo árbitro em dois momentos distintos.

A validade, é baixa (72,2%), parecendo-nos existir uma diferença de critérios de avaliação entre os peritos e os árbitros, o que se deverá conseguir melhorar com formação específica na melhoria da uniformização dos critérios de decisão.

Não encontramos diferenças significativas entre as duas categorias de árbitros, nacionais e regionais, quer para a fidelidade intra-observador quer para a fidelidade inter-observador.

No que se refere à validade, encontramos uma diferença significativa (3,4%, Sig 0,004) entre os árbitros nacionais e aos árbitros regionais, sendo esta mais alta para os primeiros.

A diferença entre o árbitro e os juízes, é significativa quer para a fidelidade intra-observador (9,9 %, Sig 0,002), quer para a validade (9,8%, Sig 0,004), sendo mais alta para os juízes.

Podemos concluir que a validade esteve sempre mais baixa do que a fidelidade intra-observador e a fidelidade inter-observador, o que denota uma divergência de critérios de avaliação entre os árbitros e os peritos.

A falta de fidelidade intra-observador parece ser a principal condicionante para um melhor desempenho.

Palavras-chave: Arbitragem, Karate, Kumite, Fidelidade, Validade, Observação, Avaliação

Abstract

An investigation centered on referees has been relatively scarce or non-existent, particularly in regards to refereeing Kumite in Karate.

The present study's objective was to characterize Karate referees' performance in a Kumite match along with its validity, as well as the intra and inter-observer reliability.

This current investigation focused on observations made by Karate referees belonging to the FPKS and entailed a Karate match that lasted two minutes. Based on that match a eight minute video was put together, where it can be seen from the perspective and position of both the referee and each one of the judges.

The data was collected during two different periods of time. The first data collected was to analyse the inter-observer's reliability and the validity. The second data collection took place within a minimum sixty day interval, in order to verify the intra-observer's reliability.

In order to analyse the validity, a group of experts was commissioned with the purpose of putting together criteria of measure.

A sample part of the video was observed individually by the referees, in a closed environment setting and documented with the use of the Match Vision Studio software.

The Bellack Index was used to assess the intra and inter-observer's reliability and validity, therefore accepting a result equal or superior to eighty percent as the guarantor of reliability (Aranha, 2007, Siedentop, 1983). The chi-squared test was used to verify if there were any significant differences in the validity between national and regional referees, as well as, in accordance with the position of the observation (referee or judge).

The reliability results were (78,7%) and (77,2%) for the intra-observer and the inter-observer reliability, respectively. The results are in close proximity to each other and one can conclude that the variability between the various referees is not greater than the variations that occur in the same referee at two different periods of time.

The validity is low (72,2%), seemingly there is a difference in the assessment criteria between the experts and the referees. This should be improved with the aid of specialised training, in order to better the uniformity of decision criteria.

In both the intra and inter-observer's reliability, no significant differences were found between the national and regional referees.

A significant difference of (3,4%, Sig 0,004) was found between the national and regional referees regarding validity, and being even higher for the national referees.

The difference between referees and judges is significant when it comes to the intra-observer's reliability which was (9,9% Sig 0,002) and (9,8%, Sig 0,004) for validity, with the judges being even higher.

A conclusion can be made that validity was always lower than the intra and inter-observer's reliability, which denotes a divergence in the assessment criteria between referees and experts.

The lack of the intra-observer's reliability seems to be a major constraining factor for a better performance.

Key words: Refereeing, Karate, Kumite, Reliability, Validity, Observation, Evaluation

General Index

Figure Index.....	viii
Table List	ix
Acronym List.....	x
CHAPTER I.....	11
1. Introduction	2
CHAPTER II	4
1. Review of Literature.....	5
1.1. Brief historic review and characterisation of the sporting discipline	5
1.2. The refereeing	8
1.3. The observation.....	13
1.4. The evaluation.....	14
1.5. Intra-observer and inter-observer reliability	16
1.6. Validity	17
1.7. The recruitment process	17
1.8. Injury prevention.....	18
1.9. Training.....	19
CHAPTER III.....	20
1. Definition of the Problem	21
2. Hypothesis Formulation	22
CHAPTER IV.....	23
1. Methodology.....	24
1.1. Sample Characterisation	25
1.2. Image Recording Protocol	26
1.3. Necessary Resources.....	28
1.3.1. Human Resources.....	28
1.3.2. Material Resources.....	29
1.4. Expert Commission Constitution.....	29
1.5. Measure Criteria Constitution.....	29
1.6. Data Collection	30
1.7. Study Variables	30

1.7.1. Independent Variables	30
1.7.2. Dependent Variables.....	30
1.8. Data treatment.....	31
1.9. Ethical and Legal Considerations	32
CHAPTER V	33
1. Presentation of results	34
1.1. Total set of marked actions	34
1.2. Set of outcome affecting actions.....	36
CHAPTER VI.....	38
1. Discussion of Results	39
1.1. Analysis of the total set of scored actions	39
1.1.1. Sample Total	39
1.1.2. Referee Categories	40
1.1.3. Observer Positioning (referee or judge)	41
1.2. Analysis of the set of result affecting actions	41
1.2.1. Sample Total.....	41
1.2.2. Referee Categories	42
1.2.3. Observer Positioning (referee or judge)	43
CHAPTER VII	44
1. Conclusions	45
1.1. Total Set of Scored Actions.....	45
1.2. Set of outcome affecting actions.....	45
1.3. Study Limitations.....	48
1.4. Recommendations.....	48
CHAPTER VIII	50
1. Bibliographic References	51
Glossary	lv
ANNEXES.....	lvii
ANNEX I.....	lviii
ANNEX II.....	lix

Figure Index

Figure 1 - The several types of kumite (Kanazawa, 2004, p. 21).....	6
Figure 2 - Disposition of human and material resources during filming.....	26
Figure 3 - Referee with mobile camera.	27

Table List

Table 1 – Criteria and categories.	28
Table 2 – Record of percentages of the total set of actions for the whole sample.....	34
Table 3 – Record of percentage of total actions for national/regional referees.....	35
Table 4 – Comparison of intra-observer reliability and validity between national and regional referees by χ^2 test	35
Table 5 – Record of the percentage of total actions for the referee/judge.....	35
Table 6 – Comparison of intra-observer reliability and validity between referees and judges using χ^2 test	36
Table 7 – Record of the percentage of point/penalty actions for the whole sample.....	36
Table 8 – Record of the percentage of point/penalty actions for the national/regional referees.....	36
Table 9 – Comparison of intra-observer reliability and validity between national and regional referees using χ^2 test.....	36
Table 10 – Record of the percentages of point/penalty actions for the referees/judges.	37
Table 11 – Comparison of intra-observer reliability and validity between referees and judges using χ^2 test.....	37

Acronym List

- FNK-P** Federação Nacional de Karate – Portugal (National Karate Federation - Portugal)
- FPKS** Federação Portuguesa de Karate Shotokan (Portuguese Shotokan Karate Federation)
- SKIF** Shotokan KarateDo International Federation
- SKIP** Shotokan KarateDo International Portugal
- WKF** World Karate Federation
- EKF** European Karate Federation
- WUKO** World Union of Karatedo Organisations

CHAPTER I

INTRODUCTION

1. Introduction

This study was developed in the scope of sportive competition refereeing in *kumite* Karate.

Considering the scarce research performed in sports refereeing, with studies performed specifically in Karate refereeing being practically inexistent, we think our research in this field to be pertinent. In this way we mean to contribute towards the scientific understanding of referee performance, orienting efforts to help improve their formation.

To understand the meaning of certain behaviours and attitudes in a training or Karate competition it is necessary to understand the context, understand the art and the culture of origin.

The referee's function is above all a guarantor of the observance of the rules defined for each event. The martial posture which characterises Karate-Do, even in its sportive aspect, should not be distorted so other results may arise in detriment of the truth and exemption in judging situations. In addition to the preparation as a referee, it is indispensable that competitors be regarded and respected as such and not as students of this or that instructor, regardless of what transpires in the competition (FPKS, 2008).

To referee a *kumite* competition in Karate, a profound technical knowledge of the sporting discipline and of the rules are required, in addition to specific training in observation. Most technical actions that lead to the scoring of points go on at high speed. As such referees must have an essential set of qualities, as described by Renon (2007), attention/concentration, capacity for abstraction, analysis and causal attribution capabilities, self-esteem and self-confidence, tolerance to criticism and frustration, empathy, communication skills, fatigue resistance, depth perception, good peripheral vision and no laterality problems, among others.

Sports *kumite* (*kyōgi kumite*) is a real combat between two opponents, directed by a team of referees in accordance to established rules, having the duration of two minutes that can be prolonged to three minutes in case of a tie, and has the sole objective of defeating the adversary.

Competition exists in both *kata* and *kumite* Karate. In this study we will specifically address the refereeing of sports *kumite*.

Participants in a Karate competition expect from the referees/judges, a fair decision-making that is according to the regulations, and which is translated in sportive truth. The way referee's decisions can sometimes negatively influence results and even put competitors' physical integrity at risk should be considered with some apprehension.

According to Guillén & Jimenez (2001), referees' performances in more or less professionalized sports, like football or basketball, may sometimes have involuntarily altered competition results. The same authors refer that the number of scientific studies about refereeing and sports refereeing is still scarce, both in the scope of psychology and in the formation/education field.

Considering that in some aspects Karate refereeing is similar to that of Judo, Escorial (2006), states that Judo refereeing is especially delicate, because a combat can be won before the end of regulation time, as the first competitor to score an *ippon* wins.

The complexity of *kumite* Karate refereeing requires an assessment of referees' performance in terms of fidelity and validity. Through the production of scientific data we aim to contribute to the improvement of refereeing quality.

So the idea of analysing the consistency of *kumite* refereeing arose, by verifying the intra-observer and inter-observer fidelity of Karate-Do referees as well as their validity. This way we hope to contribute to improve the quality of sportive truth.

CHAPTER II

REVIEW OF LITERATURE

1. Review of Literature

1.1. Brief historic review and characterisation of the sporting discipline

“Traditionally, martial arts (*budō* in Japanese), in Japan as well as Karate’s birthplace of Okinawa, began as a means of strengthening and building the mind and the body”. Kanazawa, (2006) p. 17.

As it is very difficult to know the real history of Karate, due to the lack of existing documents as well as the correction of its translations, we will start with what is called modern Karate, as we know it today which began in Okinawa, to make a small framing of Karate as a martial art and as a sport.

What is Karate then? It is a combat art characterised in its genesis as a weaponless self-defense system, using the body for that effect, in which the physical contact between opponents is controlled.

After an invitation addressed to Master Gichin Funakoshi in May of 1922, to introduce Karate in the first annual National Athletic Exhibition, endorsed by the Education Minister, in Ochanomizu-Tokyo (Figueiredo, 2006), a great diffusion of Karate began, and today it is a martial art practiced and recognized all over the world in its sportive aspect.

Karate-Do is a traditional art form of Japanese culture which became a sport of international stature, practiced around the world (Kanazawa, 1987).

According to Brito (2008) the great dissemination and public acceptance of Karate happens fundamentally for two reasons: the improvement of the individuals' physical condition and the possibility of learning a self-defense technique. It is however its sportive aspect that now begins to have more visibility.

As reported by Kanazawa (2004), today the training process in Karate encompasses four distinct areas: basic training (*kihon*), combat training (*kumite*), *kata* training (choreography of a series of movements and techniques) and competition itself; this Master, a direct student of Gichin Funakoshi, defends that competition is no longer a consequence of the traditional Karate training process. Master Kanazawa says also that current Karate can be divided into “sports Karate”, oriented exclusively for competition, and “*budō* (martial art) Karate” with a focus oriented towards discipline and training. He

furthermore states that this training process contributes not only to making an almost instantaneous reaction possible but also to the formation of excellent referees.

Its competitive aspect is divided into two different areas, as in the so called sports Karate: the *kata* competition and the *kumite* competition.

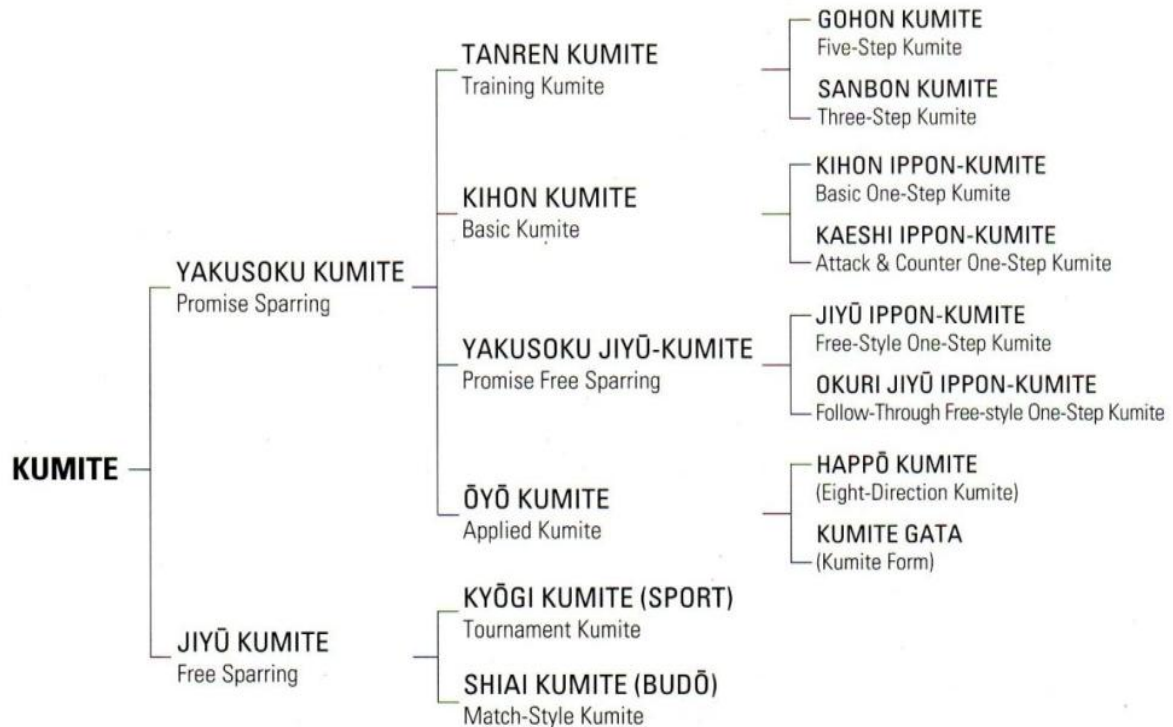


Figure 1 - The several types of kumite (Kanazawa, 2004, p. 21)

“Such a dangerous thing should not be permitted” Gichin Funakoshi in (Kanazawa, 2004, p. 11), were the words of the Master when the first *kumite* meeting began.

The first inter-university Karate training session took place on 23rd of June 1935 at the University of Keyo dojo, an orderly session of *gohon-kumite* was intended; it turned however into a totally free *kumite* session, resulting in many serious injuries. The creation of an annual inter-university meeting was nonetheless decided upon. The Universities began developing systems to make these meetings safer, and in 1955 a set of rules for the competition was created, culminating with The First All Japan Collegiate Karate Championship Tournament in 1957 (Kanazawa, 2003).

By increasing the number of specialised referees, Master Funakoshi's apprehensions in relation to *kumite* competition in Karate can be dissipated (Kanazawa, 2004).

After two international tournaments in 1968, in Los Angeles and Mexico, WUKO (World Union of KarateDo Organisations) was founded in 1970 in Tokyo, where the first Karate World Championships took place.

The first competitive event in Portugal was held in January 19th 1970 at the Colégio Salesiano do Estoril, with all combats refereed by Ronald Clark, a black belt from the Japan Karate Association (Figueiredo, 2006).

In 1972 WUKO, today designated WKF (World Karate Federation), organised the second world championship in Paris, with a Portuguese representation by the CPK (Portuguese Karate Centre).

It was also in this championship that a general unification of competitive rules was achieved. However Japanese, North American and Canadian teams boycotted the championship, refusing to compete in individual *kumite* due to problems with the team *kumite* refereeing, where there were two competition areas, a central arena and another arena closed to the public where the Japanese and Americans competed, having been defeated by either lack of contact or by excess of contact, in a refereeing deliberately against the Japanese and United States team according to the report by Master Hirokazu Kanazawa (Kanazawa, 2003).

After several World Championships, in November of 2002 the 16th World Championship, organised by WKF, took place in Madrid, and some profound alterations to the regulations were introduced, with differentiated scoring (one to three points, *Ippon*, *Waza-ari* e *Yuko*), a difference of eight points to attribute victory to a competitor, or more points after the match has ended, changing also the belt colours, now being red (*aka*) and blue (*ao*)(the latter having replaced the previous white *shiro*), probably with the intention of including Karate as a Olympic discipline and promoting its media exposure.

SKIF (Shotokan KarateDo International Federation) still maintains belt colour, red (*aka*) and white (*shiro*), as well as the *ippon* and *waza-ri* scoring system, with victory attributed to the first competitor to get *ippon-han* (in individual competition) or *ippon* (in team competition).

Ippon-han can be literally translated to “point and a half”, and can correspond to one *ippon* and one *waza-ri* or to three *waza-ri*.

As to the criteria for point attribution to a technique they are similar:

Points are given to a technique when executed to a scoring area, according to the following criteria: good form, sportive attitude, vigorous application, alertness (*zanchin*), good timing and correct distance (FNK-P, 2012).

Ippon - definition – precise and powerful *tsuki*, *ushi* and *keri* techniques, executed to the head (*jodan*), stomach, side of the torso and back (*chudan*) applying the following criteria: correct attitude and posture, abundant in spirit and *zanchin*, good distance and timing and good angle in relation to the target (ideally 90°) (FPKS, 2008).

Although great divergences always exist around international horizontal Karate organisations, rooted in the different so called “styles” or different approaches of the same culture, as demonstrated by Figueiredo (2006), today WKF, the horizontal organisation recognized by the Olympic Committee, is the main driving force behind Karate as a sport. In Portugal the National Karate Federation – Portugal (FNK-P), holder of the statute of public sports utility since 1995, represents the WKF and EKF.

And yet great vertical organisations exist, with their own identity, as is the case of Shotokan Karate-Do International Federation (SKIF), which maintains a sportive aspect of Karate closer to *budō*, including its scoring system. SKIF is implemented in over one hundred countries, integrating over two million affiliates. In Portugal it is represented by FPKS, holder of the statute of sports public utility since 1998. It is important to highlight that all FPKS instructors/trainers are registered in FNK-P, as well as competitors that wish to, depending on their trainers.

1.2. The refereeing

As there is no scientific data in the consulted bibliography regarding Karate refereeing, we strived to focus on the aspects that are common to other sporting disciplines. However, with the referee being indispensable, as its absence would warp the competitive and regulated character of the sports practice, as described by Guillén (2003), and because it is impossible to understand a sport if it lacks regulations and someone to enforce them, it is also true that the studies about sports refereeing are still scarce.

“In the opinion of trainers and competitors, referees and judges are “public enemies”. All tend to see them in a negative way. We almost always think they are against us, they will handicap us. And, as soon as they rule against us, we have a tendency to think they are persecuting us.” (Brito, 2006, p.84).

It is the competitive aspect of sports activities that determines the scope of intervention of the referee. In fact, it is during competition that we can see the technical, human and social intervention of refereeing, the extent of which cannot be undervalued as on it depends, in great part, the elevation of quality in sports practice, the defense of health, physical and moral integrity of participants in different disciplines and the reinforcement of ethical values in sports. That intervention, mid competition, is by definition personalised and is the referees' responsibility as the qualified sportive agent for that effect (Gama, et al., 2005).

Refereeing is a complex task that requires an exhaustive knowledge of the rules and good psychological capability (Gonzales-Oya & Dosil, 2007). According to Gonzales-oya and Dosil (2007), traditionally the formation of football referees is focused on an essentially theoretical background with great emphasis on the understanding of the regulations. However the technical preparation of the referee should also encompass its applications in competition. The referee trainer should disseminate knowledge of regulations and evaluate referees during competition, advising them on how their work can be more efficient. The referees' decision can be determinant in the victory or defeat of a team.

Barata-Moura (1991) cited by Gama, et al. (2005) states: The Judge is a mediator, and as such is an integrating element in the game – submitted to rules and dependant on a decisive, authorised and perceptive evaluation. The Judge-referee is the first conscience, the first reflection, of a sportive performance. He or she is, in the field and in the moment, the first assurance or referential of sportive truth of a competition.

The referees' activities are always conditioned and submitted to a set of rules that frame the competitive performances of contenders, who are in tacit agreement with said set of rules and accept them as a code for their way of competing. It is in this framework that the referees' athletic decision give validity to what is actually going on in the institutional competitive aspects of sporting practices. (Gama, et al. 2005).

The activity of a referee is extremely hard, as he must deal with several external factors, for instance: the competitors, the trainers, the public, the media, but also internal factors such as his or her qualities and physical, cognitive and even psychosocial capabilities.

In a study by Guillén, Morán and Castro (1999), cited by Guillén (2003), the main conclusions were as follows: for a large majority of referees and judges, directing a competition turns into hard or very hard work, regardless of the level of competition, be it infants or high level sports. The same study highlights the importance of the referee having been a practitioner of the sport himself, to the extent that practical knowledge of the subject favours good refereeing performance.

The monitoring of the competition, supported on the skilled observation and interpretation of the actions of protagonists and the register made by the referees' judgement, corresponding to the valuation of athletes' performances, substantiates and allows refereeing decisions, determining the validity of what transpired during the competition (Gama, et al., 2005).

The referee must "live" according to a competence primacy which is constituted of a constant demand by athletes, trainers and technicians, sports directors, social media and public (Gama, et al., 2005).

For the assessment of *kumite* sports competition in Karate, it does not suffice that a technique hits an opponent, said technique must also have: correct attitude and posture, abundant spirit and *zanshin* alertness, good distance and timing and good angle in relation to the target (ideally 90°) (FPKS, 2008).

Although the referee has to assume competitors have a moral obligation to know the rules and do everything possible not to infringe the "spirit of the rule" (Gama, et al., 2005), nowadays, according to Kanazawa (2004) there is a factor that makes *kumite* refereeing in Karate more difficult, the attempts of deceiving by competitors, which requires more and more preparation by the referees, as a great number of competitors give great importance to studying the most effective means of affecting the referees' decision.

According to Kanazawa (2004), as Karate spread across the world it also considerably changed. When the competitive system was implemented, the techniques considered of lethal force were seen as the most important. That is, to score points, a technique would, if uncontrolled, cause death or a serious injury to the opponent. Injuring contact should not occur, the technique should stop before impact, this is the representation

that referees must classify as Ippon or Wazaari. In *jiyû kumite* the several attack techniques should be controlled stopping about 3 cm from the point of impact, a practice called *sun-dome* (Kanazawa, 2004).

“The martial art actions were definitive and unique, there was no margin for mistake, as an error would cost a life and stepping back considered an act of cowardice” (Baena, 1997).

According to Filip Swennen (2009), the points scored in the current competitive model are based on a subjective evaluation of the referee and not its effective application.

Figueiredo (2006), calls it “*Karate of contradiction between the real significance of the martial technique, overcome in the ritualised regulated motion for the common benefits of both participants*”.

In Guillen's (2003) opinion, in the first place we should know the problems that affect refereeing and sports judging in general and in each of the different sporting disciplines, subsequently creating training programs.

The technical training of a referee is a decisive factor in sporting quality, and in the interventions they must carry out during the individual and collective performances of competitors, and can be divided into three aspects: technical, human, and social (Gama, et al., 2005).

The referee's technical training allows him to respond positively to the competence demands that are made, and to be considered the guarantor of the competitions' validity. This training, according to Gama, et al., (2005), should focus on three domains: rules, refereeing techniques and regulations corresponding to the level of competition. Guillén (2003) also makes reference to the referee's technical competence, to the degree of knowledge of the regulation and the ability to apply it objectively, as well as to the training and experience.

On the human aspect the referee needs to have a strong personality, which is created by virtue of the manifestation of his or her honesty, integrity, impartiality, decorum, righteousness, technical rigour and public spirit (Gama, et al., 2005).

Regarding the social aspect, refereeing puts on a significant scope, because of the appreciation caused by the influence it can have in the elevation of the technical and human quality in sports practice, defending health, moral and physical integrity of participants of different disciplines, and reinforcing the ethical values of the sport. (Gama, et al., 2005).

According to Gama, et al., (2005) a referee should present certain characteristics such as good physical condition, good psycho-physiological condition, good emotional condition, and good technical conditioning. These should be intimately related and dependant on the sportive and pedagogic characteristics of their sport. The same authors also define the main attributes of the referee's technical profile as: visual and auditory acuity, immediate perception of situations and decision opportunity in judging.

These characteristics fit into what Guillén (2003) construes a referee must possess; an ensemble of physical and psychological capabilities adequate to a good performance of its function. Of all the factors affecting refereeing the most important and transversal to all sports are probably the psychological factors.

According to Guillén (2003) the main psychological function implied in refereeing is decision making. Deciding is an action resulting of the interaction of different activities like understanding and interpreting. The making of a decision is based in the sensory response system which constitutes our physiological support (seeing, hearing and touching) and allows understanding. This physiological support cannot be confused with psychological function, but is a necessary condition for it.

Decision making is not only knowledge of the rules or lack there of, but a complex mental process that can be improved, enhancing the competence of those judging a competition. It is a choice or judgement opted for in detriment of other alternatives. It is made based on information which is constantly mutating in the context of the competition (Gama et al., 2005).

In the referee's performance there are naturally a variety of decision, derived from human nature and inherent to human characteristics. However this variation should not put sportive truth at stake.

Also Tuero et al. (2002) state that decision making, understood as the capacity to make decisions quickly, giving the impression of having been absolutely certain of what was seen, is one of the valuable qualities a good referee should have. Other psychological qualities of a good referee are: self-esteem, referees should trust themselves; concentration ability, referees should know which stimuli and relevant situations they should attend to, and which they should ignore; anxiety is also an important factor to take into account, anxiety can lead to a reduction in concentration and wrong decision making.

The qualities cited by Tuero et al. (2002), are fundamental, and can grant a sensibility appropriate to refereeing, but only those able to add to it the technical attributes that only experience, career and training confer will ever truly be referees.

1.3. The observation

For Gama et al. (2005), learning to observe is one of the main tasks of the referee, so it is essential that, beyond the domain of the regulations that rule the sport, he is worried about constantly improving his competence in observation.

The referee's main mission is to make observations for evaluation, according to the rules and principles, which should have as an orienting model the knowledge of the situations to evaluate (Gama, et al. 2005).

The swiftness, rhythm and alternation of real competitive situations, defined in the scope of each discipline's specificity, extremely hinder the absolute correctness of referees' decisions (Gama, et al. 2005).

According to Piasenta (2002), seeing at great velocity requires habituation, the beginner observer of an action at first apprehends only the epiphenomena, consequences, effects while the cause, at real speed, eludes his sight. The variation of observation, in direction, velocity and duration, depends, among other factors, on what reaches for and captures the eye, and so continuous formation of referees in applied observational methodology seems to us indispensable.

In the understanding of Gama, et al. (2005), the technical formation of the referee has as necessary condition the acquisition of a seeing know-how and a reading know-how, directed at the competitive situations that are characteristic of each sporting discipline. Seeing know-how relates to the "employment" of vision, with the knowledge and experiences from the specific discipline. Therefore if the referee has practised it, as a competitor or trainer, he will potentially have better seeing and reading know-how.

To have reading know-how means to have the capability to identify and interpret the signs that characterise the scoring of points and the infractions or fouls committed by competitors, which is extremely important, as the referee should only intervene, by interrupting, when a competitor scores a point, when illicit actions occur or immediately after the respective occurrence.

According to Anguera (1990), observing depends on the interaction of three fundamental elements: perception, interpretation and previous knowledge.

The referee's activity utilizes in the first instance the visual perception of the athlete's movements, and through this "mental footage" he must take definitive decisions in a few seconds. Furthermore, in some sporting disciplines, the observation of sportive movements is also associated with technical analysis, which requires the referee's observational task to be even more scrupulous. Anguera (1997), cited by Sarmiento (2004), states that observing can be an easy or a hard task, depending on the data observed, and giving utmost importance to the formation of qualified observers. However elaborated the observation plan is, it will only be effective according to the observers' abilities (Sarmiento, 2004).

For Gama, et al., (2005), the observation of an action registered on video, indirect observation, can provide the referee a comparison point with analysis carried out by other referees and verify possible differences in evaluation. Nonetheless to watch a video register is not the same as seeing an event live.

1.4. The evaluation

The referee is constantly evaluating the competitors actions, that is to say comparing those actions with the information he has about rules and evaluation parameters apprehended in the referee's formation. He or she observes, gathers information, processes it and has to decide accordingly.

To evaluate is to estimate, to appreciate, to calculate something's value. In this sense, evaluation is a constant human activity, as at any given moment we must collect information from the environment, evaluate that information and make decisions accordingly. It is a basic information processing mechanism for human beings.

The information received from the environment is compared to the information available in memory structures, allowing for judgement on the value of that information and determining subsequent action (Rosado & Colaço, 2002).

For Sánchez (2006) to evaluate consists in assigning a value judgement to something or someone. In this sense, evaluation is a common activity and can constitute actions, such as collecting information, emitting judgement based on comparison and making a decision.

A referee can know intimately the sportive discipline and the rules of the competition, but that is not sufficient for him or her to have a good performance. Also of importance are the employment of the sense organs and the processing of information in the brain.

Sometimes the referee must measure the competitors' intentions, which cannot be seen, but inferred from the interpretation of actions compared to the corresponding regulation (Guillén, 2003).

Frequently there is some confusion between the terms measuring and evaluating. Measuring means indicating performance, is a quantitative description, while an evaluation comprises of both the quantitative description and a qualitative description, in addition it can formulate value judgements. However evaluation may or may not be based on measurements (Sánchez, 2006).

To know a sports discipline inside out is to have a deep and concrete understanding of the discipline's rules, not only what they say, but also what is implicit and what they allow us to presuppose (Gama, et al., 2005).

Therefore the referee must know the regulation and be competent in its application, first learning the rules and then learning to apply them.

According to Caracuel, Andreu & Pérez (1995), besides the regulation, one should take into account the domains to which the referee/judge is refereeing the perceived facts, such domains depend on their interpretation of regulation as well as other factors (previous experiences, conduct styles, motivation, situation etc.)

Man stores the information gathered by the senses, compares it to previous experience and uses it or not in responsive interactions with the world (Caldas, 2000).

The referee must have a refined sense of sight as well as perceptive selection (attention), this plays several roles in perception. One of them is preparation: we are better prepared to perceive when we pay attention. Attention also allows us to select certain aspects of a scene to analyse, ignoring others (Gleitman, Fridlund & Reisberg, 2003).

As perception is a cognitive process, it is necessary to take information to the mind before we can do something with it. During perception, knowledge about the world is combined with constructive abilities, physiology and the experiences of the perception subject (Davidoff, 2001).

Still, in humans, the main way of physically selecting the input of stimuli is eye movement. Several studies show that subjects direct their look more frequently to the most

visually informative areas. The pattern of gaze fixation may vary for different observers (Gleitman, 1999).

Pizzera & Raab (2012), who studied the relation between previous visual and motor past experiences and the current performance of referees in football, handball, ice hockey and trampolines, have concluded that cognitive judgements are related to the referees' previous experiences as athletes or spectators.

According to Hancock (2011), in a study performed about the perception of ice hockey referees, comparing experienced and novice referees, results showed that experienced referees position themselves better, prioritize visual behaviours, seemingly learning from experience and making better use of peripheral vision. It concludes that in referee training focus should be given to visualisation training with video games, using this context to help referees learn to direct their visual attention.

1.5. Intra-observer and inter-observer reliability

A referee must be consistent in his or her decisions, that is, facing the same situation in different moments similar decisions will be made.

According to Gama et al. (2005), athletes and trainers expect the referees to be consistent, i.e. their decisions to be the same facing the same or similar circumstances.

Rodrigues (1997) states that intra-observer reliability intends to guarantee that the same observer in different moments makes a similar codification of the several studied behaviours.

If necessary, intra-observer reliability can be established, comparing the results of the observation of the same video in another occasion (Thomas & Nelson, 1996).

In view of the same situation, different referees should decide in a similar fashion, the inter-observer reliability can be established in some qualitative research studies with the evaluation of a competent agent (Thomas & Nelson, 1996).

In the act of observation, it is necessary that those refereeing a sporting event see it the same way. Every time an action is observed, facts are noticed to which one should add the criteria for measuring and evaluation of observed facts (Guillén, 2003).

Individuality acts as a constraint, actively and dynamically, throughout the observation process, so a totally homogeneous intra-observer and inter-observer tuning

cannot be expected, but instead a relative tuning, which increases proportionally to the experience in observation and/or specific observation training (Sarmiento, 1991).

1.6. Validity

It is, without a doubt, the decision of the referee that determines the validity of competitive performances, defining sportive truth. This fact forces us to consider the referee's decision as final, and if the referee applied the regulations correctly, then we can say the decision is valid.

The principle of validity reports to the fact that the instruments used should measure what they propose to (Tenbrick, 1999).

To determine observation validity of the sample, results obtained were compared with the measure criteria defined by the experts commission.

1.7. The recruitment process

“The natural recruiting process for referees should focus on elements of the sportive discipline – ex-athletes, ex-players or trainers, who, having had a concrete experience of the discipline, would opt for a career in refereeing” (Gama, et al., 2005) p. 40-41

In Karate the recruitment process referred to above has been a reality so far, it even occurs that most referees are trainers, and were or are still competitors, as the last grade of Karate competition, veterans, starts at age 70.

The athletes' grade is a relevant pre-requisite for the function of Trainer in all questioned countries, as well as for Refereeing Technicians...

The function of a Refereeing Technician is classically assumed by higher ranking trainers. In the first Karate tournament in Portugal in 1970 the referee was Ronald Clark, exactly the technician who officiated the vocation of the pioneers of the first competitive participation.

The function of the Refereeing Technician, both in the role of judge in the Kata discipline, as well as in the roles of judge and referee in the Kumite discipline, is the natural evolution for the Trainer and evaluator of his or her disciples. It is a role identical to that of the graduation judge, which is also always assumed by the wisest trainer; the master. In that sense, it is natural that the most mature Karate trainer, and some of the pioneers, assume the functions of refereeing technicians (Figueiredo, 2006, p. 551).

For Caracuel, Andreu & Pérez (1995) to be a good referee it is not indispensable to have been an athlete of the sport, yet it always has its influence, but depending on the experiences he had as a contestant, this influence can be positive or negative.

For Gama, et al. (2005), the referee's task is in a way similar to that of the trainer, if we analyse it. Both need to know the technique in the movements, both need to know how to observe well, and differ only in their intervention method: the trainer teaches/corrects, the referee evaluates.

Therefore, we always have the problem of preparation, training and update. A referee has the need to have specific training for his function. As discussed above the referee recruitment process is focused on the most experienced instructors/trainers, who have a great technical knowledge of the discipline and its regulations, specific training is fundamental however.

1.8. Injury prevention

One of the objectives of sporting activities should be health maintenance, so the referees' performance is important in the domain of injury prevention. According to Critchley, Mannino and Meredith (1999), in a study where the objective was to document the injury rate in three consecutive Shotokan Karate championships in England in 1996, 1997, and 1998, where protections were forbidden and control mandatory, it was concluded that a strict refereeing is essential to maintain control and minimise contact.

Arriza, Leyes, Zaeimkohan, and Arriaza (2009) state that the most important factors in reducing injury rate have been the disposition for referees to stop the competition when it becomes dangerously rough, penalising contestants that disregard safety in competition, forbidding certain uncontrolled attack methods and strictly applying the existing rules for the competition.

The regulations regarding excessive contact, which causes injuries, are clear as to the penalties to apply, however the evaluation criteria still have large undefined areas (Swennen, 2009)

1.9. Training

A referee's training is permanent. His laboratory the competition he directs daily. The intervention of further training structures must be permanent.

The more organic training, associated to career definition and its category, should be intimately connected to its practical framing. The referee's training must be fundamentally directed towards the resolution of concrete situations (Gama, et al. 2005, p. 41-41).

As described by various authors, the decision-making process is fundamental to the performance of referees. The use of video could be an instrument for their training. A video recording, with segments of actions which require decision, can be created and presented to the referee, improving the decision-making process in sporting contexts (Gama et al., 2005).

The preparation of referees is not limited to the attendance of the initial training courses which enable them to practice the activity. They need a system of continuous preparation considering that, unlike competitors, they cannot train in competition-like conditions (Gama et al., 2005).

Improving the observational, analytical and decision-making competences can be done through discussion and training with more experienced referees. This discussion allows the referee to acquire knowledge on how to resolve a particular situation (Gama et al., 2005).

One of the techniques which is already part of the training plan of athletes and is recently recognized by investigators as an optimisation factor for their performance, is imagery. This concerns the mentally reliving or creating of experiences. Imagery is possible because the human mind can recall past events and recreate them as images. According to Gama, et al. (2005) our mind is capable of reconstructing what we lived in the past differently, according to our intentions. This mental repetition allows us to deal with difficult situations in a more efficient manner, facing them more confidently.

According to Catteeuw, et al. (2010), in a study over two training formats designed to improve decision-making over offsides in football, video simulation and computer animation, concluded that cognitive/perception ability training in both formats had a positive effect on decision making, and therefore should be taken into consideration towards referee/judge training.

CHAPTER III

DEFINITION OF THE PROBLEM AND FORMULATION OF HYPOTHESIS

1. Definition of the Problem

The objective of refereeing in sports Karate competition, as in other sporting disciplines, is to achieve the so-called sportive truth, that is, the referee makes decisions according to the rules.

According to Gama et al., (2005) it is the referee's decision which determines the validity of sportive performances. It is the only way to elevate the quality of sports practice, to defend the health and physical as well as moral integrity of competitors and to reinforce ethical sporting values.

For that objective to be met intra-observer reliability is necessary, which means the referee rules in a similar fashion in similar situations. And that all referees will rule the same way facing the same situation so that inter-observer reliability also exists.

It is the objective of this research to characterise the performance of Karate referees in *kumite* competition with regards to its validity, that is, the intra-observer and inter-observer reliability.

The validity of refereeing is based on the referee's decision making, supported according to the regulation and supported in the seeing and reading know-how applied to the competitors' actions (Gama, et al., 2005).

The referee, before making a decision, is an observer, and according to the methodology of the observation systems, should possess a degree of intra-observer reliability, which aims to guarantee that the same observer, in different moments, presents the same results.

The referee should also possess a certain degree of inter-observer reliability, which assures that different observers using the same observation system make decisions similarly.

We face therefore the problem of studying reliability and validity of Karate referees in *kumite* (combat) sportive competitions.

In pursuance of our objectives, we identified the following problem:

Is it that Karate referees'/judges' decisions are valid, that is, do they translate sportive truth?

2. Hypothesis Formulation

To be able to answer our problem, the following research hypothesis were formulated:

- There are significant differences in intra-observer reliability, inter-observer reliability and validity between national and regional referee categories in *kumite* Karate.
- There are significant differences in intra-observer reliability, inter-observer reliability and validity in *kumite* Karate refereeing related to the position of the observer (referee or judge).

If we can identify in this study some of the problems that affect refereeing, judging and the consequent decision by the referee/judge, we may contribute to answer the most direct and urgent needs of refereeing through specific training programs, so referees improve their performance.

CHAPTER IV
METHODOLOGY

1. Methodology

The comparison between refereeing systems is not the object of this work; still we had to opt for a system, in this case, the present study was performed according to the FPKS/SKIF rules regarding scoring system, but it can be replicated using FNK-P/WKF.

In both systems the refereeing panel for each combat consists of a referee (*shushin*), four judges (*fukushin*), and a combat supervisor (*kansa*).

While it is true that only the points *ippon* and *waza-ri*, and the penalties that result in *ippon* or *waza-ri* for the adversary, can lead to victory or defeat, we decided to answer the question under investigation in two ways:

1- Utilising the categories in a more comprehensive manner (competitor, level, technique, point, warning/penalty and others), as the referee can influence the result, by not seeing an action he was supposed to (*mienai*), or attributing an *ippon* or *waza-ri* to a technique that lacked correct attitude and posture, abundant spirit and *zanshin*, good distance and timing and a good angle in relation to the target (ideally 90°).

Therefore we start by analysing intra and inter-observer reliability and validity for a total set of scored actions, as defined for the recording (competitor, level, technique, points, warning/penalty and others), that is to say, all frames where experts considered an action would lead to stopping the combat for a decision by the referee panel, or when any element of the sample asked the video to be stopped and registered an action.

2- Utilising only the categories that influenced results (competitor, points and warnings/penalties), as in practice they can influence the result of a *kumite* event and lead to sportive truth.

So we will only analyse the intra and inter-observer reliability and validity for the set of actions which influenced the result of the combat (competitor, points and warnings), as this set of actions influences sportive truth and can effectively respond to the problem raised.

1.1. Sample Characterisation

Of the universe of 36 referees/judges of the Portuguese Shotokan Karate Federation, a sample of 21 referees/judges was conveniently constituted, according to their availability on the dates when data collection was scheduled, and divided into two categories, national referees (age 45 ± 10 min. 32, max. 62, refereeing experience 15 ± 8 min. 3, max. 27) and regional referees (age 34 ± 8 min. 22, max. 50, refereeing experience 3 ± 2 min. 1, max. 5).

It is worth referring that the referees/judges are practitioners, are or were competitors and the majority, 81% of the sample, are trainers.

The present investigation was based on the observation, by FPKS Karate referees, of a Karate combat with a two minute duration. A video was produced of that combat with a duration of eight minutes, which reproduced the combat observed from the referee's position, as well as from the position of each of the judges. For that effect five digital cameras were used, one of which mobile, assembled on the referee's head, and four stationary cameras in the judges' positions (Figure:2). The combat was performed by two experienced senior competitors, both black belt athletes, one of whom was national champion for FPKS in 2012.

Given its purpose, and so as to not influence the sample referees' decision, it was opted, against FPKS rules, that there would be no interruption by the referee in the two minutes of combat.

Three distinct videos were edited, after which a commission of experts selected one of them as being the one presenting the most difficult decision-making situations.

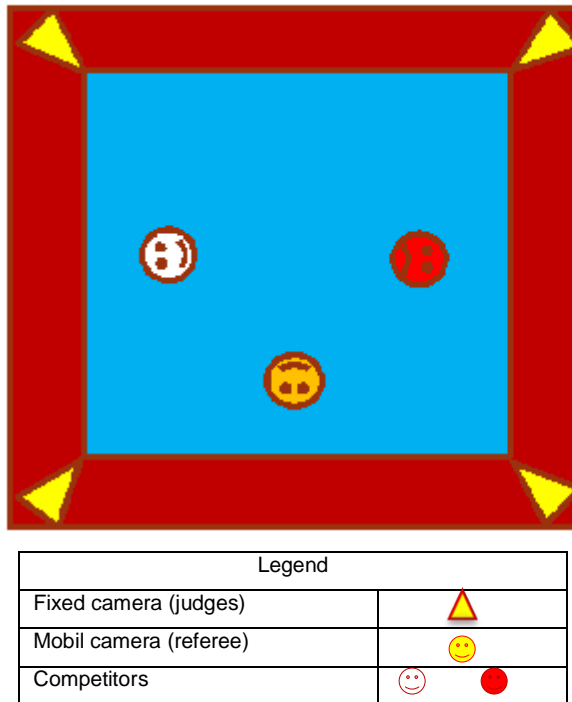


Figure 2 - Disposition of human and material resources during filming.

For the making of the video six competitors, four camera operators and one experienced international level referee were contacted. A competition area of 8 x 8m was set up, where the three combats took place.

1.2. Image Recording Protocol

- All the material was assembled and tested before the beginning of the combats.
- The stationary video cameras were placed on tripods outside the competition area, just as the judges, being operated and directed at the competition area.
- The capture of images inside the competition area was achieved with a mobile video camera assembled on the head of an internationally experienced referee who was positioned in what would be the best angle of vision.



Figure 3 - Referee with mobile camera.

Although all the material was tested, one of the cameras failed to capture parts of the combat, and so its results were discarded; there was no possibility of repeating the image capture under the same conditions. It was considered as not compromising the results, as the three other stationary cameras captured the combat. Furthermore in a *kumite* competition there may not be enough judges, so it can take place with one referee and two judges in stationary positions or even two referees in mobile positions.

The set of images captured by the four cameras was edited into a single eight minutes film, so that the combat could be analysed from the various angles (Referee and Judges).

After collecting the video recordings, the content was transferred to a PC, converting it to digital format, using the software Magix Terra Tec ® and DVD Vídeo Soft FreeStudio. The viewing of the recordings was done on PC using Match Vision Studio Premium® software.

Table 1 – Criteria and categories.

Criterion	Category	Description
Competitors	<i>Aka</i>	Competitor with red belt.
	<i>Shiro</i>	Competitor with white belt.
Level	<i>Jodan</i>	Technique executed at head level.
	<i>Chudan</i>	Technique executed at torso and abdominal level.
Technique	<i>Tsuki</i>	Attack technique with fists.
	<i>Keri</i>	Attack technique with feet.
Points	<i>Ippon</i>	Precise and powerful techniques of <i>tsuki, ushi e keri</i> , executed towards the head (<i>jodan</i>), stomach, side of the torso and back (<i>chudan</i>).
	<i>Waza-ari</i>	Technique not with the extent of <i>Ippon</i> , means a bit more than 50% of the <i>Ippon</i> .
Warning/ Penalty	<i>Mubobi</i>	Warning for lack of care, the competitor didn't protect himself.
	<i>Jogai</i>	Out of bounds, left contest area.
	<i>Mukirioku</i>	Lack of martial spirit.
	<i>Keikoku</i>	Warning for dangerous action.
	<i>Hansoku chui</i>	Warning for excessive action, <i>waza-ari</i> for the opponent.
	<i>Hansoku</i>	Expulsion of the competitor who committed an excessive action to the opponent.
Others	<i>Aiuchi</i>	Simultaneous technique.
	<i>Mienai</i>	The referee did not see the technique.
	<i>Yowai</i>	Weak technique, without form or <i>zanshin</i> .

In table 1 we can find the criterion and categories created in the Match Vision Studio Premium® software to register the actions, according to the regulation and following the criteria for *kumite* refereeing.

1.3. Necessary Resources

1.3.1. Human Resources

- 4 Experts;
- 6 Competitors;
- 1 Referee;
- 4 Camera Operators;

1.3.2. Material Resources

As far as material resources the following were necessary:

- 1 Laptop Computer;
- Magic Movies on DVD Terra-Tec Software;
- Freestudio video conversion Software;
- Match Vision Studio Premium® Software;
- PAWS statistical Software (version 18);
- 4 Video Cameras, with incorporated microphone;
- 4 Tripods for the video cameras;
- 1 mobile video camera. Gopro;
- 1 Chronometer;
- Strips to outline the competition area;

1.4. Expert Commission Constitution

For the constitution of the expert commission the following criteria were used: they should all be active international Shotokan Karatedo International Federation (SKIF) referees, with a minimum grade of 6° Dan. The expert commission was constituted of four specialists: Mário Águas Shihan, 7° Dan, Nobuaki Kanazawa Sensei, 6° Dan, Luís Filipe Ferreira Sensei, 6° Dan, José Mário Mil-Homens Sensei, 6° Dan.

1.5. Measure Criteria Constitution

The videos were visualised by the experts commission, gathered in a closed room, with the purpose of obtaining a measure criteria which corresponded to perfect refereeing of the combat, having chosen the video with the most situations which are usually harder to decide for the refereeing team.

The experts visualised the video together, as many times as they saw fit, at normal speed and in slow-motion, until reaching a consensus over the actions liable to be intervened by the referee and judges, deciding over what transpired in the combat.

A measure criteria was obtained by registering the time unit and frame of all actions liable of registry.

1.6. Data Collection

After the video was validated by the expert commission, the referees were contacted individually to watch it.

The data collection was done in two moments, the first aimed to analyse validity and inter-observer reliability, and the second, with an interval of at least sixty days, to verify intra-observer reliability.

The video was viewed by each member of the sample individually, in a closed environment, using the Match Vision Studio software with which the registry of the observation was made. Because there can be a difference between the time at which the viewer requested a pause and the frame where the action took place, all recordings of actions were corrected to the frame in which the referred action occurred.

Before the first viewing, each referee was shown a video of another two minute long combat so as to adapt them to the situation, as there are differences between a real combat and the viewing of its recording.

It was explained to the referees how the data was to be collected, that is to say, that they would watch the video at normal speed, with no chance for repetitions, refereeing the combat and asking to stop the video every time they thought a relevant action occurred, evaluating it like they would in a real combat, where they would make decisions regarding scoring of points, penalties or nothing (*tomarisen*).

1.7. Study Variables

1.7.1. Independent Variables

- Referee categories – National or Regional
- Observation Position – Referee or Judge

1.7.2. Dependent Variables

- Referee's decisions

- Intra-observer reliability – Comparison between each referees first and second observation;
- Inter-observer reliability – Comparison between different referees' decisions;
- Decision Validity – Comparison between referee/judge decisions and the criteria measure defined by the experts;

1.8. Data treatment

In order to verify the consistency of the referees' evaluation, the results of the first and second viewing were compared, thus determining intra-observer reliability; to do so the results of the *frame by frame* action record were compared, if different the result would be 0, if the same the result would be 1, in the end the mean score was found.

To determine inter-observer reliability the results of the observation of the referees in the first data collection were compared, allowing us to verify how the referees evaluated, we compared the *frame by frame* record between referees, if different the result was 0 if the same the result was 1, in the end we calculated the mean score.

To determine validity, the results of the first observation were compared to the measure criteria defined by the experts (appendix I).

The record of actions was analysed *frame by frame*, and action by action, that is, each set of actions registered by the referee in each frame was compared to the set of actions (measure criteria) defined by the experts.

We used Bellack's index, commonly used to measure fidelity, and which formula consists of “(total agreements/total agreements + disagreements) x 100.” (Aranha, 2007, Siedentop, 1983). In a larger sample 85% is an acceptable value, in a small sample (i. e., n = 30) a result of 80% or more is accepted as a guarantor of reliability.

The descriptive analysis and testing of inter and intra-observer reliability (Bellack) were made using PAWS (version 18) statistical software.

To verify if there were significant differences in validity between national and regional referees and between referees and judges a chi-squared test was used, and a significance of $p \leq 0,05$ adopted for the two hypothesis.

1.9. Ethical and Legal Considerations

To perform this work, authorisation was requested of the President of the Portuguese Shotokan-Karate Federation and of the President of its Refereeing Council to contact the referees for them to participate in the study.

To all the questions formulated in the data collection and treatment, anonymity of the intervening subjects was guaranteed, with the interests of the institution being placed above the interest of the observers.

CHAPTER V

PRESENTATION OF RESULTS

1. Presentation of results

In this chapter we present the results according to the same order followed in the methodology section, that is, analysing first intra-observer reliability, then inter-observer reliability and finally validity. First we present the analysis pertaining to the total of the sample, followed by analysis according to referee category and observation position (judge or referee). This methodology will be followed by both the analysis of the set of total scored actions and the set of actions that resulted in the scoring of points

Annex I shows the record of the experts decision we will use as our measure criteria. The expert panel registered 18 *frames* where the combat would be interrupted to make a decision; of those 18 moments, 8 would lead to the scoring of points or penalties that could influence the result of the viewed match; of the other 10 *frames*, in 4 of them the attacks did not meet the norm of what the regulation defines as a scoring point (*yowai*), in the remaining 6 there seems to have been a valid attack, but from the referee/judge's angle of view it is not possible to score a point (*mienai*).

Of the records made by the sample set, 43 *frames* (annex II) were registered where a judge/referee scored an action. As can be verified, the number of scored actions is around three times that of the measure criteria, leading us to believe that at least some of the judges/referees, despite identical training in evaluation criteria, identified more situations requiring the combat to stop, scoring actions, penalties and other actions than those that actually occurred.

1.1. Total set of marked actions

In Table 2, we can verify intra and inter-observer reliability and validity for the total sample, having as reference the total set of marked actions.

Table 2 – Record of percentages of the total set of actions for the whole sample.

	Sample
Intra-observer Reliability	68,7%
Inter-observer Reliability	69,0%
Validity	64,1%

Table 3 shows us the difference between intra and inter-observer reliability and validity in the two referee category analysed, national and regional referees.

It is in the intra-observer reliability that the largest difference between referee categories can be found.

Table 3 – Record of percentage of total actions for national/regional referees

	National	DP	Regional	DP
Intra-observer Reliability	70,1%	7,87	66,3%	10,75
Inter-observer Reliability	70,3%		72,7%	-
Validity	64,4%	12,01	63,7%	10,09

The difference isn't significant, as we can see in table 4. As to the inter-observer reliability, it is higher for regional referees than for national referees.

Table 4 – Comparison of intra-observer reliability and validity between national and regional referees by χ^2 test

	Intra-observer Reliability			Validity		
	Value	df	Sig.	Value	df	Sig.
Pearson Chi-Square	1,46	1	,22	,05	1	,82

Table 5 shows us the difference between referees and judges. When compared, both regarding intra-observer reliability and validity, results were higher for the judges.

Table 5 – Record of the percentage of total actions for the referee/judge

	Referees	SD	Judges	SD
Intra-observer Reliability	65,8%	14,34	69,6%	8,95
Inter-observer Reliability	70,6%		68,4%	-
Validity	63,6%	16,74	64,3%	10,30

Also in this regard, the difference is not significant, as we can see in table 6, with inter-observer reliability amongst referees found to be higher for referees than for judges.

Table 6 – Comparison of intra-observer reliability and validity between referees and judges using χ^2 test

	Intra-observer Reliability			Validity		
	Value	df	Sig.	Value	df	Sig.
Pearson Chi-Square	1,17	1	,27	,03	1	,85

1.2. Set of outcome affecting actions

We now present the results for intra and inter-observer reliability and validity only for the set of actions which define the result of the combat (competitor, points and warnings), as this is the set of actions that influences sportive truth, and which can effectively answer our problem.

Table 7 – Record of the percentage of point/penalty actions for the whole sample.

	Sample	SD
Intra-observer Reliability	78,7%	10,08
Inter-observer Reliability	77,2%	-
Validity	72,2%	12,87

As we can verify in table 8, intra-observer reliability is similar in both referee categories, and inter-observer reliability is higher in the regional referees category, with significant differences between the two categories, as can be seen in table 9.

Table 8 – Record of the percentage of point/penalty actions for the national/regional referees.

	National	SD	Regional	SD
Intra-observer Reliability	78,7%	9,20	78,8%	12,06
Inter-observer Reliability	77,8%		81,6%	-
Validity	73,5%	14,72	70,1%	9,64

Table 9 – Comparison of intra-observer reliability and validity between national and regional referees using χ^2 test.

	Intra-observer Reliability			Validity		
	Value	df	Sig.	Value	df	Sig.
Pearson Chi-Square	,001	1	,981	8,171	1	,004

Table 10 shows us the difference between the evaluation of referees and judges, with an increment of about 10%, both on intra-observer reliability and validity, and a rise of inter-observer reliability when subjects are placed in the judge position.

Table 10 – Record of the percentages of point/penalty actions for the referees/judges.

	Referee	SD	Judges	SD
Intra-observer Reliability	71,4%	17,08	81,3%	9,11
Inter-observer Reliability	73,2%		78,6%	-
Validity	64,9%	16,32	74,7%	12,45

In this case the difference is significant, both in the intra-observer reliability, as we can see in table 11, as well as in validity, which seems to indicate increased difficulty for the referees' position in compared to the judges' position.

Table 11 – Comparison of intra-observer reliability and validity between referees and judges using χ^2 test.

	Intra-observer Reliability			Validity	
	Value	df	Sig.	Value	df
Pearson Chi-Square	9,905	1	,002	8,171	1

CHAPTER VI

DISCUSSION OF RESULTS

1. Discussion of Results

In the present chapter the reflection over the data resulting from the study is presented. The purpose of this reflection is to answer the question under investigation.

In *kumite* refereeing, there are a number of situations that can lead the referee to interrupt the combat, without any penalties or points being scored, ex. (weak attack, lack of *hiki-te* or *hiki-ashi*, lack of *kime* or *kiai*, lack of *zanshin* or *maai*). However the judges should always give the referee some kind of indication as to the action that motivated the interruption.

For this reason differentiated analysis for the total set of scored actions and the set of point scoring actions were performed.

1.1. Analysis of the total set of scored actions

Next we will analyse the record of the total set of observed actions (competitor, level, technique, points, warnings and others). Any incongruity is considered as lack of reliability or validity depending on the analysis performed.

1.1.1. Sample Total

As far as intra-observer reliability, the average observed in the sample is below the accordance value defined by literature, 80% according to the Bellack Index (Aranha, 2007, Siedentop, 1983), which means each of the referees facing the same situation viewed with a minimal time interval of 60 days, as referred in the methodology chapter, has a variation of around 11.3% below that minimum criteria and 31,3% below what would be total consistency, showing us a high individual fluctuation, translated in personal inconsistency in decision-making.

The average observed for intra-observer reliability is similar to that for inter-observer reliability, meaning that an improvement should be sought by enhancing individual consistency. After solving this issue the accordance between referees should subsequently be looked at.

According to Sarmiento (1991) the individuality is a constraint that actively and dynamically acts during the process of observation, and as such a totally homogeneous intra and inter-observer fine-tuning cannot be expected, rather a relative fine-tuning, that increases proportionally to the experience in observation and/or observation specific training.

Validity, the comparison between the average standard deviation in the sample and the measure criteria defined by experts, is even lower (between 4 and 5%) than those of reliability, which was predictable, as a high individual fluctuation, and an evaluation criteria of the same order between referees, makes it hard to be in agreement with the measure criteria. This can mean some situations are viewed by the experts in one way, and in another by most referees.

1.1.2. Referee Categories

The largest observed difference between the two categories of referees was in intra-observer reliability, 3,8% higher in the national referees, although not a significant one.

As to inter-observer reliability, it is higher in the regional referee category, with a difference of 2,4%.

Validity was higher in the national referee category by less than 1%, with no significant difference between the two categories.

The results are below expectations and the non-existence of significant differences in the analysed parameters regarding referee categories, leads to the hypothesis that the origin of the problem is the constant subjectivity of the regulations.

Also the so called lack of martial posture (FPKS, 2008), that is, the disrespect for the norms of conduct, the adversary, exacerbated manifestations after the scoring of a point, commentaries or discordant gestures towards a decision from the judging panel, can lead to penalties like the annulment of an *ippon* or *waza-ari*.

The competitor's intention must sometimes also be measured (Guillén, 2003). The evaluation of the martial posture should infer the actions of competitors, comparing them to what is defined as an evaluation criteria in the corresponding regulation.

However, if a regulation makes reference to, for example, correct form and vigorous application, “A technique is said to have correct form when its characteristics confer a probable efficiency within the conceptual parameters of traditional *karate*.”

The vigorous application defines the potency and speed of a technique, as well as the palpable will of it succeeding” (FPKS, 2008; FNK-P, 2012). These criteria have a high level of subjectivity, which makes *kumite* refereeing extremely difficult.

1.1.3. Observer Positioning (referee or judge)

When we compare the results according to the position of the observer as referee or judge, both intra-observer reliability, as well as validity are higher for the judge position, although the difference is not significant, with inter-observer reliability higher for the referees' position than the judges'.

Although theoretically the referee should always be better positioned to observe, the observation is harder, as there is movement, both of the observer (referee), as well as the observed object (competitors), contrary to judges in their fixed positions.

When comparing the standard deviation in relation to position of observation and total sample, the observation position does not seem to influence the performance parameters of referees.

1.2. Analysis of the set of result affecting actions

We now analyse the intra and inter-observer reliability and validity for the set of actions that defined the result of the combat (competitor, points and warnings), as this is the set of actions that influences sportive truth and effectively answer our problem.

1.2.1. Sample Total

The percentage of agreement rise about 10%, in relation to the total set of actions bringing it close to the 80% minimum, according to the Bellack Index, which is natural, as there are fewer number of actions to evaluate.

In relation to intra-observer reliability, although it reached 78,7% it is not yet a guarantee of reliability (Aranha, 2007, Siedentop, 1983), showing an elevated individual fluctuation. According to Rodrigues (1997) the intra-observer reliability aims to ensure that the same observer in different moments makes a similar codification, this way we think, the low intra-observer reliability is translated in personal inconsistency in decision-making,

as Guillén (2003) states the main psychological function implied in refereeing is the decision-making, resulting from different activities, such as understanding and interpreting.

Intra-observer reliability reached 77,2%, a value that does not ensure reliability, showing us that there is a difference in the evaluation criteria that does not permit for all the referees to evaluate the same way. However, if the problem of individual fluctuation is solved, so will the issue with inter-observer reliability, as the values are very close, which tells us that the referees evaluate, generally in the same way, both between themselves and in different situations.

As to validity, it is still low, 72,2%, although the difference if compared to the inter-observer reliability in the total set of actions analysis dropped slightly, with a bigger difference in evaluation criteria between experts and referees than between referees, which in a way corroborates what Filip Swennen (2009) says, the points scored in the present competitive model are based on the referee's subjective analysis.

The fact that validity is lower than inter-observer reliability, can mean a drift in the refereeing style of most referees from that recommended by the experts.

When analysing the final result of the combat, we verified that its result would be a draw (*hikiwake*), in the experts decision, while the result of the sample's refereeing would give the victory to the competitor in white belt (*shiro*), reinforcing the idea of a need for further training that leads to a standardisation of evaluation criteria.

However, analysing the *frame* where the experts scored a point to the red belt competitor (*aka*), we verified that the technique was executed at high speed, anticipating an attack from the adversary, which made it difficult to see, requiring a high level of attention/concentration.

It seems to us that improving referee performance can be achieved with specific training, aimed at making evaluation criteria standard.

1.2.2. Referee Categories

In intra-observer reliability a difference cannot be found between the two categories of referees. We find a similar value in both categories as well as the total sample.

As to inter-observer reliability, it is higher for regional referees than for national referees, showing that regional referees has a bigger uniformity in evaluation criteria.

Regarding the validity of national referees, the difference is higher in relation to regional referees, which was predictable given their wider experience. This is a significant

difference, in line with Hancock (2011), who comparing experienced and novice ice-hockey referees, verified that they seem to have learned from experience.

Focusing on what is essential (points), national and regional referees display significantly different performances, justifying the classification, which seems adjusted. The fact that these differences do not occur when the total set of actions is analysed could mean that the national referees are more focused on aspects that are more relevant to the combat outcome.

1.2.3. Observer Positioning (referee or judge)

When we compare the results according to the observation position as referee or judge, the intra-observer reliability as well as the inter-observer reliability and validity in the judge position are higher, with a significant difference for the intra-observer reliability as well as for validity.

In the judge position, both the intra-observer reliability as well as validity rise about 10%. It seems to us that refereeing is harder from the referees' position than from the judges' position, as far as we understand because of the movement of the referee, which can lead to perception errors. It is also possible that the experience most referees have as a judge, compared to the experience they have as a referee, conditions their performance, because of the familiarity with the observation position.

CHAPTER VII
CONCLUSIONS

1. Conclusions

1.1. Total Set of Scored Actions

The results obtained allow us to verify that when analysing a comprehensive set of actions that can lead to an interruption in the combat, although some of them don't have an influence on its outcome, the levels of intra and inter-observer reliability as well as validity for the total sample set, are below the accepted values in consulted literature (80%).

Analysing the values of intra-observer reliability reveals a high individual fluctuation, which is translated in personal inconsistencies in decision-making, influencing inter-observer reliability and consequently validity.

When comparing the two referee categories, national and regional, the biggest difference between them, albeit not significant, is in intra-observer reliability, with higher results for national referees, who show more consistency. As to inter-observer reliability, it is higher for the regional referees, revealing a better standardisation in evaluation criteria, which unfortunately leans away from sportive truth, as confirmed by the higher validity for national referees.

In the analysis of results based on observation position of either the referee (mobile position) or the judges (fixed positions), both intra-observer reliability and validity are higher for judges, although the difference is not significant, with inter-observer reliability being higher for referees than for judges. These results show that observation by the referee is more difficult, as there is movement by the observer, unlike the judges who stay in fixed positions.

1.2. Set of outcome affecting actions

For the set of actions that influence sportive truth and which effectively answers our problem, we can verify that the accordance values are close to the minimum value of 80%, not being however a guarantee for reliability.

Regarding intra-observer and inter-observer the values are very close between referees and judges, which tells us that in general they evaluate the same way in similar situations.

As to validity, although still low, it seems to us that there is a difference in evaluation criteria between experts and referees, which should be improved with further training, specifically in improving decision criteria. As evaluation is based on partially subjective criteria a good effort of interpretation of the wording and spirit of the regulation is necessary.

When comparing the two categories of referees, national and regional, we can see there are no significant differences between the two groups, with regards to intra-observer and inter-observer reliability. As far as validity, the national referees increased the difference compared to regional referees, which was predictable given their larger experience. This was a significant difference. The fact that this analysis diverges from that where the total set of actions is considered may indicate that the national referees' actions are more focused on the more relevant aspects of the combat, while the regional referees' actions are dispersed in less relevant aspects.

The difference in positioning of referee and judges is significant, be it for intra-observer reliability or for validity, which indicates the task presents more difficulty to the referee than to the judges, in our understanding because of the referees' movement, which can lead to perception errors and consequently to decision-making mistakes.

Also the inter-observer reliability is higher for judges than for referees. However the lack of statistic significance in this difference does not allow us to consider this a tendency.

Validity, which is the guarantee for sportive truth, was always lower than intra-observer and inter-observer reliability, denoting divergent evaluation criteria between referees and experts, that cannot be explained by lack of reliability.

Performance also seems to be conditioned by the observation position. The results suggest it is harder to referee from the referee position than from the judges position.

Verification of the formulated hypotheses

In order to solve our problem we had presented the following hypotheses:

- There are significant differences regarding intra-observer reliability, inter-observer reliability and validity between regional and national category referees in *kumite* Karate refereeing.

We partially accept this hypothesis, when considering only the actions that score points or penalties, and which thus contribute to the outcome of the combat. In these instances the national referees present a significantly superior level of agreement with the experts.

- There are significant differences regarding intra-observer reliability, inter-observer reliability and validity in *kumite* Karate refereeing depending on the position of the observer (referee or judges).

We partially accept this hypothesis, when considering only the actions that score points or penalties, and which thus contribute to the outcome of the combat. In these instances the referees presented a higher intra-observer reliability and a better level of agreement with the experts when positioned as judges than when in the referee position.

As no other quantitative analysis studies were found in the scope of sports refereeing to compare our results with, we were limited to comparing our results to the limit described in literature as the guarantee for intra and inter-observer reliability and of validity.

We can conclude that the performance of the FPKS Karate referees is characterised by lack of intra-observer reliability, that is, we don't seem to be in the presence of different ways of refereeing, but rather the main problem appears to reside in the lack of consistency of each referee. Facing the same situation, the same referee should make a similar decision. This lack of individual consistency could be related to a deficiency in standard criteria or specific training, or even the inherent subjectivity of the regulation.

As there are no big differences between intra-observer reliability and the various parameters, we can deduce this to be the main cause for lack of validity.

The lack of intra-observer reliability appears to meet several authors descriptions, especially in studies performed in the scope of sports psychology, where refereeing is considered a extremely difficult activity, given the multiplicity of factors involved, both of the external order, as well as the referees' internal factors.

There is, however, about 5% difference between reliability and validity, this could be due to different refereeing styles. These questions can and should be researched, but the criteria fluctuation of each referee is now evident and should be the first issue to be corrected.

1.3. Study Limitations

Taking into account future research into this subject, we found as a limitation to the study the fact that it cannot be done in real context, as the combat had to be performed and managed by the competitors, without interruptions by the referee.

The referee that recorded the combat with the mobile camera is an international referee, with more than 20 years experience, so he should have placed himself in the best possible position, which may not be the case with a less experienced referee.

The viewing of a video is different from reality, which can influence the interpretations of the actions by the referees.

In this study each referee had to decide individually over each action. However, in a real combat the decision is taken by the referee in accordance with his opinion and that of the judges.

In this study there was no influence of external factors such as: the trainers, the public and the other combats happening simultaneously, which would happen in a competitive context.

1.4. Recommendations

As stated in the literature revision, the referee's activity is extremely difficult. The referee must deal with several internal factors, such as their qualities and physical, cognitive and even psychosocial capabilities. The referee's actions will naturally have a variation in decision, resulting from human nature and inherent to the characteristics of humans, which should not however put sportive truth at stake.

Considering the conclusion of the study, and having identified herein some of the problems that affect refereeing within the FPKS, we recommend referee training actions, based on the lack of intra-observer reliability, making the referees aware of this issue, offering the specific training, for instance, the viewing of a video in the beginning and end of their training program and confronting the trainees with their results.

Also in the referee training certain aspects related to the psychology of refereeing should be taken into account, honing their capabilities in this scope, improving their psychological qualities.

Based on the differences of results for validity, we also find important that the training programs provide observation specific training and especially in evaluation criteria standardisation.

We suggest the present study be replicated in light of the new FNK-P regulation, so as to permit the generalisation of results and eventual adoption of corrective measures to vertical or horizontal organisations.

CHAPTER VIII

BIBLIOGRAPHIC REFERENCES

1. Bibliographic References

- Anguera, M. (1990). Metodología observacional. In J. Arnau, M.T. Anguera & J.Gómez. *Metodología de la investigación en Ciencias del Comportamiento*. Murcia: Secretariado de Publicaciones de la Universidad de Murcia, 125-236
- Aranha, A. (2007). *Observação de aulas de Educação Física*. Vila Real: Universidade de Trás-os-Montes e Alto Douro.
- Arriaza R.,Leyes M., Zaeimkohan H. & Arriaza A., (2009). The injury profile of Karate World Championships: new rules, less injuries. *Knee Surg Sports Traumatol Arthrosc* (2009) 17:1437–1442.
- Brito, A. V. (2008). A Influência da Variabilidade do Objectivo Final da Realização da Acção Motora na Coordenação Neuromuscular de Movimentos Balísticos Realizados com o Membro Superior. Tese de doutoramento apresentada à Universidade Técnica de Lisboa, Faculdade de Motricidade Humana, Lisboa.
- Brito, A.P. (2006). *Iniciação à Psicologia do Desporto para Atletas*. Gráfica 2000.
- Caldas, A. C. (2000). *A Herança De Franz Joseph Gall, O Cérebro ao Serviço do Comportamento Humano*. Alfragide: McGRAW-HILL.
- Caracuel, J.C., Andreu, R. & Pérez, E. (1995). Análisis psicológico del arbitraje y juicio deportivos. Una aproximación desde el modelo interconductual. *Motricidad*, 1, 5-24.
- Catteeuw, P. Gilis, B.Jaspers, A. Wagemans, J. & Helsen, W. (2010). Training of Perceptual-Cognitive Skills in Offside Decision Making. *Journal of Sport & Exercise Psychology*, 32, 845-861.
- Critchley G.R., Mannion S. & Meredith C.,(1999). Injury rates in Shotokan Karate, *Br J Sports Med*, 33:174–177.

- Davidoff, L. L. (2001). *Introdução à Psicologia*, 3ª ed. S.Paulo: Pearson Education do Brasil Ltd.,
- Escorial, M. S. (2006). Características deseables para el arbitraje y el juicio deportivo en Judo, *Revista Iberoamericana de Psicología del Ejercicio y el Deporte*, 1, 27-40.
- Figueiredo, A., (2006). *A Institucionalização do Karaté, Os Modelos Organizacionais do Karaté em Portugal*. Tese de doutoramento apresentada à Universidade Técnica de Lisboa, Faculdade de Motricidade Humana, Lisboa.
- FNK-P, (2012). Regras de competição, <http://www.fnkp.pt/pages/24>.
- FPKS, (2008). Regulamento de arbitragem.
- Gama, A., Corte, A. Monge, D. M., Araújo, D., Ramos, J. P., Horta, L., Marivoet, S., Serpa, S., Weinberg, R. & Richardson, P. (2005). *Manual do Árbitro*, Centro de Estudos e Formação Desportiva, 2ª ed. Lisboa: Instituto do Desporto de Portugal.
- Gleitman, H., (1999). *Psicologia*, 4ª ed. Lisboa: Fundação Calouste Gulbenkian.
- Gleitman, H., Fridlund A. J. & Reisberg, D. (2003). *Psicologia*, 6ª ed. Lisboa: Fundação Calouste Gulbenkian.
- Gonzales-Oya, J. & Dosil, J. (2007). *La Psicología Del Árbitro de Fútbol*, Noia: Touxosoutos, S.L.,
- Guillén, F. & Giménez, H. (2001). Características deseables en el arbitraje y juicio deportivo, *Revista de Psicología del Deporte*, 10, núm. 1, 23-34.
- Guillén, F. (Coord) (2003), *Psicología Del Arbitraje e el Juicio Deportivo*. Barcelona: INDE Publicaciones.

- Hancock, D. (2011), *Examining Perceptual Differences Amongst Elite, Intermediate, and Novice Ice Hockey Referees: Visual Attention and Eye Movement Recordings*. Thesis submitted to the Faculty of Graduate and Postdoctoral Studies In partial fulfillment of the requirements For the PhD degree in Human Kinetics, Faculty of Health Sciences School of Human Kinetics. University of Ottawa.
- Kanazawa, H. (1987). *S.K.I. Kumite Kyōan*. Tokyo: Autor.
- Kanazawa, H. (2003). *Karate My Life*. Singapore: Kendo World Publications Ltd.
- Kanazawa, H. (2004). *Karate Fighting Techniques* (1^aed.) Tokyo: Kodansha International, Ltd.
- Kanazawa, H. (2006). *Black Belt Karate The Intensive Course* (1^aed.). Tokyo: Kodansha International, Ltd.
- Piasenta J. (2002). *Aprender a observar* (ed. Portuguesa). Lisboa: Centro de Estudos e Formação Desportiva.
- Pizzera, A. & Raab M. (2012). Perceptual Judgments of Sports Officials are Influenced by their Motor and Visual Experience, *Journal of Applied Sport Psychology*, 24: 59–70. Association for Applied Sport Psychology.
- Rosado, A. & Colaço, C. (2002). Avaliação das Aprendizagens, fundamentos e aplicações no domínio das actividades físicas. Lisboa: Omniserviços,
- Renom, J. P. & Halcón, A. P., (2007). Jueces y oficiales de regatas a vela; motivaciones, percepciones y formación, *Revista de Psicología del Deporte*. Vol. 16, núm. 1 pp. 55-66, Universitat de les Illes Balears, Universitat Autònoma de Barcelona.
- Rodrigues, J. (1997). Os Treinadores de Sucesso, Estudo da Influência do Objectivo dos Treinos e do Nível de Prática dos Atletas na Actividade Pedagógica do Treinador de Voleibol. Lisboa: Ed. FMH.

- Sánchez, D.B. (2006). *Evaluar en educación física*. Barcelona: 9ª ed. Inde.
- Sarmento, P. (1991). Observação na formação. *Revista Horizonte*, VII, 41, 167-174.
- Sarmento, P. (2004). *Pedagogia do Desporto e Observação*. Lisboa: Ed FMH,
- Siedentop, D. (1983). *Developing Teaching Skills in Physical Education*. Mountain View: Mayfield Publishing Company
- Swennen, F. (2009). The Evolution of Karate: From Secret Martial Art to Worldwide Cultural Sport. <http://www.eskk.co.uk/uploads/2/4/3/3/2433362/thesisbudai.pdf>.
- Tenbrink, T. (1999). Construir listas de control y escalas de evaluación. In T. Tenbrink (Ed.), *Evaluación: Guia Práctica para Docentes* (5ª Ed. pp. 257-273). Madrid: Narcea Ediciones.
- Thomas J.R. & Nelson J.K. (1996). *Research Methods in Physical Activity* (3ª ed.). S. Paulo: Human Kinetics.
- Torres Baena, F. (1997). El karate deportivo, estudio praxiológico; análisis del kumite deportivo en la modalidad de shiai-kumite : (estudio histórico y kinantropométrico). Tese de Doutoramento apresentada à Universidade de Las Palmas de Gran Canaria, Departamento de Edudación Física, Las Palmas.
- Tuero, C. Tabernero, B. Guillén, F. & Márquez, S. (2002). Análisis de los Factores que Influyen en la Práctica del Arbitraje. *Revista Scape*, vol. 1, núm 1, pp. 7-16.

Glossary

Aiuchi	Simultaneous
Aka	Red
Chudan	Body's middle level (chest, abdomen)
Fukushin	Referee
Gohon-kumite	Combat set in five steps, where the attacker warns as to the level and kind of attack
Hiki-ashi	Returning or pulling foot
Hiki-te	Returning or pulling hand
Hiki-wake	Draw
Ippon	Point
Jiyu kumite	Free sparring
Jodan	Upper level of the body (face area)
Jogai	Leaving the competition area
Kansa	Kumite supervisor, normally a more experience referee
Kata	Pre-arranged forms comprising a series of offensive and defensive techniques that are performed individually against one ore many imaginary adversaries.
Keri	Kick (pronounced “ <i>geri</i> ” when attached to another word, as in <i>mae-geri</i>)
Kiai	Loud vocalization timed to coincide with a decisive technique, combative spirit.
Kihon-ippou kumite	Kumite set in one steps, where the attacker warns as to the level and kind of attack.

Kime	Focus
Kumite	Sparring
Kyōgi kumite	Tournament kumite according to the rules of the competition
Maai	Distance maintained between opponents during kumite
Mienai	Did not see
Sambon-kumite	Kumite set in three steps, where the attacker warns as to the level and kind of attack.
Shiai kumite	(<i>Budō</i>) Free combat with rules dictated only by the personal conscience of each opponent in full control of attacks
Shiro	White
Sun-dome	Bringing attacks to a sudden halt just before making contact with the target (one “sun”, unit of measure, sensibly 3 cm)
Sushin	Judge
Torimasen	Nothing
Tsuki	Punch (pronounced “ <i>zuki</i> ” and also written like so when connected to another word, ex. <i>gyaku-zuki</i>)
Uchi	Strike
Waza-ari	Half a point
Yowai	Weak attack, lacking form
Zanshin	A state of physical and mental preparedness following an attack.

ANNEXES

ANNEX I

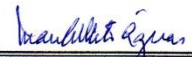
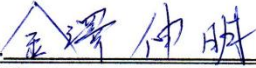
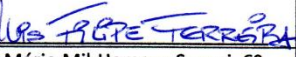

Registo de validação da observação do vídeo de kumite pela comissão de peritos

Validation register of the video kumite observation by the expert committee

Critério Criterion	Time Frames	Tempo Seg Time Sec	Competido res Competitor	Nível Level	Técnica Technic	Pontos Points	Penalizações Penalties	Outros Another one
Categoria Category			Aka Shiro	Jodan Chudan	Zuki Geri	Waza-ari Ippon	Muboubi Jogai Keikoku Hansoku-chui Hansoku	Aiuchi Mienai Yowai
	627	0:00:25	0	0	0	0	0	Yowai
	1224	0:00:50	0	0	0	0	0	Yowai
	1950	0:01:19	0	0	0	0	0	Yowai
	2572	0:01:45	Aka	0	0	0	Jogai	0
	2800	0:01:54	Shiro	Chudan	Zuki	Waza-ari	0	Mienai
	3712	0:02:32	0	0	0	0	0	Yowai
	4432	0:03:01	0	0	0	0	0	Mienai
	5195	0:03:33	0	0	0	0	0	Mienai
	5646	0:03:51	Aka	0	0	0	Jogai	0
	5881	0:04:01	0	0	0	0	0	Mienai
	7520	0:05:08	Aka	Jodan	Zuki	Waza-ari	0	0
	7713	0:05:16	0	0	0	0	0	Mienai
	8109	0:05:32	0	0	0	0	0	Mienai
	8733	0:05:58	Aka	0	0	0	Jogai	0
	8968	0:06:07	Shiro	Chudan	Zuki	Waza-ari	0	0
	10587	0:07:14	Aka	Jodan	Zuki	Waza-ari	0	0
	11799	0:08:03	Aka	0	0	0	Jogai	0
	12035	0:08:13	Shiro	Chudan	Zuki	Waza-ari	0	0

Azambuja-Portugal, 2011, October 22

A Comissão de Peritos, The expert committee

Mário Águas Shihan, 7º Dan Ass. 	Presidente e Diretor Técnico da Federação Portuguesa de Karate Shotokan (FPKS) Árbitro Internacional, Shotokan Karate-Do International Federation, (SKIF)
Nobuaki Kanazawa Sensei, 6º Dan Sign. 	Kancho Designate, Educated Taisho Univ. SKI World Champion In 2000, SKI All Japan Champion in 1996-2000, Japan National Champion in 2004, Metropolitan Champion 3 times, JFK Champion 6 times. International Referee, SKIF.
Luís Filipe Ferreira Sensei, 6º Dan Ass. 	Sub-Diretor Técnico da FPKS, Árbitro Internacional SKIF.
José Mário Mil-Homens Sensei, 6º Dan Ass. 	Presidente do Conselho de arbitragem da FPKS, Sub-Diretor Técnico da FPKS, Árbitro Internacional, SKIF.

ANNEX II

Record Made by the total sample

<i>Frame</i>	<i>Time Seconds</i>	<i>Competitor</i>	<i>Level</i>	<i>Technique</i>	<i>Points</i>	<i>Warnings/Penalties</i>
450	0:00:18					
627	0:00:25	0	0	0	0	0
1017	0:00:41					
1224	0:00:50	0	0	0	0	0
1542	0:01:03					
1950	0:01:19	0	0	0	0	0
2124	0:01:27					
2325	0:01:35					
2572	0:01:45	Aka	0	0	0	Jogai
2800	0:01:54	0	0	0	0	0
3249	0:02:13					
3516	0:02:24					
3712	0:02:32	0	0	0	0	0
4306	0:02:56					
4432	0:03:01	0	0	0	0	0
4624	0:03:09					
4861	0:03:19					
5028	0:03:26					
5195	0:03:33	0	0	0	0	0
5407	0:03:41					
5646	0:03:51	Aka	0	0	0	Jogai
5881	0:04:01	0	0	0	0	0
6335	0:04:19					
6599	0:04:30					
6794	0:04:38					
7061	0:04:49					
7520	0:05:08	Aka	Jodan	Tsuki	Waza-ari	0
7713	0:05:16	0	0	0	0	0
8109	0:05:32	0	0	0	0	0
8287	0:05:39					
8733	0:05:58	Aka	0	0	0	Jogai
8968	0:06:07	Shiro	Chudan	Tsuki	Waza-ari	0
9403	0:06:25					
9667	0:06:36					
9863	0:06:44					
10131	0:06:55					
10587	0:07:14	Aka	Jodan	Tsuki	Waza-ari	0
10777	0:07:21					
11015	0:07:31					
11353	0:07:45					
11560	0:07:54					
11799	0:08:03	Aka	0	0	0	Jogai
12035	0:08:13	Shiro	Chudan	Tsuki	Waza-ari	0