Original Article

A stakeholder assessment of basketball player evaluation metrics

JOSE A. MARTÍNEZ , LAURA MARTÍNEZ

Departamento de Economía de la Empresa, Universidad Politécnica de Cartagena, Cartagena, Spain

ABSTRACT

Martínez JA, Martínez L. A stakeholder assessment of basketball player evaluation metrics. *J. Hum. Sport Exerc.* Vol. 6, No. 1, pp. 153-183, 2011. In this research we examined the opinions of basketball stakeholders regarding several questions of special interests to valuate players. Players, coaches, agents, journalists, editors, bloggers, researchers, analysts, fans and chairs participated in this macro-research. After analysing their opinions using the content analysis methodology, we found that current player evaluation systems are insufficient to fulfill the expectations of stakeholders regarding the definition of value, because they fail to rate intangibles. In addition, the importance of qualitative thinking is prominent and should be considered in valuating such intangibles. The current system of valuation used in Euroleague and Spanish ACB League (Ranking) is acknowledged as deficient, but stakeholders think that other advanced metrics do not significantly outperform Ranking. Implications for management, decision making and marketing in basketball are finally discussed. **Key words**: PLAYER EVALUATION METRICS, BASKETBALL, STAKEHOLDERS, DECISION MAKING.

Corresponding author. Departamento de Economía de la Empresa, Universidad Politécnica de Cartagena, Cartagena, Spain.
E-mail: josean.martinez@upct.es
Submitted for publication November 2010.
Accepted for publication March 2011.
JOURNAL OF HUMAN SPORT & EXERCISE ISSN 1988-5202
© Faculty of Education. University of Alicante
doi:10.4100/jhse.2011.61.17

INTRODUCTION

The statistical revolution will not be televised[2] This is the slogan of the Association for Professional Basketball Research (APBR), one of the most important places to develop, interchange and discuss cutting edge research on the application of statistics to address basketball research problems. This is only one of the multiple forums dedicated to sports analytics emerged from the "Moneyball" phenomenon (Lewis, 2003), which cover academic-professional research (e.g. the Journal of Quantitative Analysis in Sports, and the Sloan Sports Analytics Conference), and non-academic research (see Berri and Bradbury, 2010, for a discussion about the scientific contributions of both worlds).

In the specific field of basketball, the APBR site (www.apbrmetrics.com) is the main forum to disseminate new contributions to basketball analytics, where major researchers, professional and non-professional analysts constantly participate. One of the most recurrent topics for the APBR community is the evaluation of basketball players, i.e. how quantitative analysis may help to more "objectively" valuate players. The importance of this topic is prominent, and has been a matter of subject of articles in academic journals (e.g. Berri, 1999; Berri and Bradbury, 2010; Cooper, Ruiz and Sirvent, 2009; Esteller-Moré and Eres-García, 2002; Piette, Annand and Zang 2010), seminal books (e.g. Berri, Schmidt and Brook, 2006; Oliver, 2004; Winston, 2009), prospective books (e.g. Hollinger, 2005; Doolittle and Pelton, 2009), and a plethora of specialized websites (e.g. www.basketball-reference.com, www.82games.com, www.basketballvalue.com). The performance/ability/efficiency/value of players on the court, are terms used to evaluate and translate player performance into a number. This information is valuable for many disciplines, such as economics (e.g. measuring the productivity of players), marketing (e.g. measuring the market value of players), management (e.g. decision to sign players), or training (e.g. searching the best combination of players on the court). It is an interesting topic for the broad discipline of sports, and specifically for sport management.

Martinez (2010a) reviews more than 200 systems for evaluating player performance in basketball. Some of them are complementary, but many of them are competing methods which differ in terms of theory and methodology. While contradictory views are vital in a scientific discipline, the current situation is not desirable, because the dissemination and use of this information impacts economic, marketing, managerial or training decisions. For example, Martínez and Martínez (2010d) and Cooper, Ruiz and Sirvent, (2009) proposed two new methods to evaluate players, based on a probabilistic approach and on data envelopment analysis, respectively. The application of these methods to the ACB (Asociación de Clubes de Baloncesto) Spanish League changed the rankings of players in several statistical categories. Awards such as the Most Valuable Player (MVP), and other recognitions derived from being the leader in categories such as points, rebounds, assists, etc. are a function of the method that has been applied to compute the final score of players in each category. This could influence the media value of players, because leaders in each category increase their exposure to media, and hence their market value. Another example of this bias is the use of "Ranking", "Efficiency" or "Net Plus Minus" as official systems to evaluate players in the main competitions of the world (e.g. NBA, Euroleague, ACB League). Martínez (2010b), in his review of 83 leagues around the five continents, shows. that these three systems (Ranking is a little variation of Efficiency) are widely criticized by academic and professional experts (e.g. Berri and Bradbury, 2010; Winston, 2009). Decisions made based on these systems may be erroneous but still, these systems continue to be used. Due to a plethora of systems, existing information increases exponentially, thus

[2]*This slogan is based on the song of Gil Scott-Heron, an American poet and musician. The song was writing in the seventy's, and it was a critic about people alienated by television.*

complicating decision-making. Decisions based on multiple options can be inefficient as compared to decisions made on a few considerations. This is a recurrent topic that has been studied in psychology, marketing and economics (e.g. Ariely, 2008; Gigerenzer, Todd and ABC Research Group, 1999; Kuksov and Villas-Boas, 2010).

The status quo is maintained because of a lack of a deep understanding of several key questions that form the basis of the rating systems. Questions such as how to define the value of a player, the role of qualitative evaluations in quantitative analysis or the proper methods to evaluate player value must be clarified before implementing any method to evaluate player performance.

We use a stakeholder assessment framework to address these questions. We analyze the opinions of a panel of experts composed of players, coaches, agents, journalists, analysts, researchers, bloggers and fans regarding several questions about the following issues: (1) How to define the value of a player; (2) The role of qualitative evaluations in quantitative analysis; (3) The adequacy of the system of evaluation of Euroleague and ACB Spanish League; (4) The adequacy of the MVP selection system of the ACB Spanish League; (5) The relationship between player evaluation metrics and market value; (6) The knowledge, preferences and use of some of the most widely used evaluation systems; and (7) The best methods to assess the value of a player.

We analyze the opinions of a sample of these stakeholders in order to depict the current view of the different interest groups. The aim of this research is not to force consensus, but to understand the current thinking of stakeholders regarding these fundamental questions. However, consensus may be desirable in order to reach an agreement on critical questions[3]. Consensus may be the starting point for developing more accurate rating systems to evaluate players.

The stakeholder approach is a noticeable paradigm in strategic management (Doyle, 2002). It takes into account the opinion of the different groups which interact with the company. In this research, the core problem is the player evaluation metrics, and stakeholders have different levels of appreciation. Therefore, players are evaluated, coaches use these metrics to make decisions, journalists write their columns in function of the rating system used, agents bargain about contracts, fans elaborate opinions, etc. The stakeholder paradigm provides an extensive analysis to address problems concerning many groups.

This paper shows that current player evaluation systems are insufficient to fulfill the expectations of stakeholders regarding the definition of value, because they fail to rate intangibles. In addition, the importance of qualitative thinking is prominent and should be considered in valuating such intangibles. The current system of valuation used in Euroleague and Spanish ACB League (Ranking) is acknowledged as deficient, but stakeholders think that other advanced metrics do not significantly outperform Ranking. However, the poor knowledge of other advanced stats by the vast majority of stakeholders limits their uselfuness. Researchers and analysts are more critical of the Ranking system. If more efficient managerial decisions in basketball must be made, it is crucial to consider what stakeholders think about rating systems.

[3]*Achieving consensus through expert's judgements is a widely used approach. This is used in disparate arenas such as medicine or linguistics, among others. For example, Colditz, Atwood, Emmons, Monson, Willett, Trichopoulos, and Hunter (2000) used group consensus among researchers at the Harvard Medical School and Harvard School of Public Health to identify risk factors as definite, probable, and possible causes of cancer, then they created the Harvard Cancer Risk Index. Likewise, in linguistics, consensus among members of the Royal Academy of Language of Spain is needed to include a new linguistic term in the dictionary.*

 © 2011 University of Alicante

Research questions

This research addresses seven topics, which are of special interests to the basketball analytics field.

(1) The definition of the "value of a player"

Establishing the definition of the construct to be studied is the first step to measure it. This is a core principle in science, especially in social sciences (Hayduk, 1996), where constructs may have different meaning to disparate people (Fransella, Bell and Bannister, 2004). Beyond philosophical discussions about the nature of reality, there is agreement among critical realism and relativism about the necessity to construct a system of representations collectively shared. Human perceptions are these representations (Barfield, 1988), and are the basis for the social construction of reality. This is the idea behind the concept of intersubjectivity; there is no objective reality, but knowledge and facts are what society agrees to be (Berger and Luckman, 1966). This is also the position adopted by the foundationalist perspective (Thompson, 1990), in an effort to integrate positivism and interpretivism trough humanism. It is the search of truth through consensus[4], and it is the basis over people construct their reality (Fransella, Bell and Bannister, 2004), shared meanings, beliefs and metaphors (John, Locken, Kim and Monga, 2006; Zaltman, 2003).

Systems to valuate basketball players are heterogeneous in terms of what they are measuring. Some of them measure efficiency in a relative way (the final value of a player depends on the efficiency standard of other players) but do not consider inputs (Cooper, Ruiz and Sirvent, 2009), other measure value as a linear combination of statistical categories (see Martínez, 2010a, for a review), but there is heterogeneity regarding which categories have to be used, and their weights. Some systems link value to team wins (Berri, 1999), and others link value to the performance of the team when the player is on the court (Winston, 2009). A clear definition of value is necessary to implement systems to measure it.

(2) The role of qualitative evaluations

Quantitative analysis of sport is achieved to get a more "objective" view of the reality of the games, and to avoid the bias of subjective perceptions of stakeholders. This is a form to break with some social conventions, and with the consistency of cognition inherent to human beings (Ariely, 2008), a recurrent problem to incorporate new ideas to society, which have been considered as an important constraint to the dissemination of scientific knowledge in economy, sociology or psychology (Berri, Schmidt and Brook, 2006; Levitt and Dubner, 2005).

However quantitative analysis has limitations. Lewin and Rosenbaum (2007) found that sophisticated metrics do not outperform simple forms to explain the salary of players and team wins in the NBA[5]. They concluded that sometimes the subjective judgement of experts may be most useful than any numeric system to evaluate player performance. A related finding is showed by O'Donoghue and Williams (2004). They compared predictions about results of the Rugby World Cup of 2003 achieved by a panel of experts, and predictions resulted from disparate statistical methods, using different algorithms and assumptions. Quantitative analysis did not significantly outperform human judgments. A similar result was found regarding this comparison among "human" and "machine" predictions in the Soccer World Cup of 2002. Others like Lewis (2003), and Skinner (2010) demonstrate the ability of quantitative analysis to yield insightful results, which would be very difficult to reach relying only on human judgement (see also Berri

[4]*It is also true that some criticism to this line of reasoning is done by philosophers such as Thomas Kuhn or Paul Feyerabend (Rothman, Greenland and Lash, 2008).*

[5]*This is an unpublished work written by these two well-kown analysts. This work has been extensively criticized by David Berri in his blog, and discussed in the APBRmetrics forum.*

and Schmidt, 2010). In addition, many of the most successful NBA teams count on statistical analysts (Doolittle and Pelton, 2009), because they are convinced that quantitative analysis yields competitive advantage.

But, what do stakeholders think about qualitative evaluations to valuate basketball players? Recall that the vast majority of MVP awards granted in all the competitions around the world totally of partially depends on qualitative evaluations of experts (media, coaches, etc.) and, to a less extent, vote of fans (Martínez 2010b). In addition, some mass media provide qualitative evaluations of player performance after each game. Sometimes, there are players with "bad numbers" that are qualitatively better evaluated than other players with "great numbers". Sometimes, the best player of the game selected by journalists does not match with the player with the highest quantitative performance. In Spain, important media such as Marca, As or Gigantes del Basket, provide such qualitative evaluations.

(3) The adequacy of the "Ranking" system.
The "Ranking" system is one of the most used methods to valuate players in official competitions around the world. As Martínez (2010b) shows, it is used in the national leagues of Mexico, Croatia, Slovenia, Spain, Italy, FYROM or Israel, and in the following supranational leagues: Adriatic, Baltic Eurocup and Euroleague. Euroleague and Spanish ACB League are the two most important competitions outside the NBA.

This system valuates players in function of some statistics derived from the box-score. Specifically, it is simply a linear combination of positive actions (points + rebounds + steals + assists + blocks + fouls drawn) minus negative actions (missed shots + turnovers + fouls + block against). Other systems such as "Efficiency" are very similar (it does not consider fouls drawn, fouls and block against). Efficiency is also very popular around the world, and it is used in the national leagues of Brazil, USA (NBA), Malaysia, Germany, Austria, Bulgaria, Cyprus, Denmark, Finland, France, Greece, Holland, Island, Latvia, Luxemburg, Poland, Romania, Switzerland and Turkey. In addition it is used in the following supranational leagues: Balkan and Eurochallenge. Other minor modifications of "Ranking" and "Efficiency" are used in Argentina, Uruguay, Hungary, Slovakia, Lebanon or Czech Republic. Spearman correlations among all these similar systems to valuate players are above 0.99 (Martínez, 2010c).

Berri and Bradburi (2010) or Winston (2009) show the critical flaws of "Efficiency" as a system to valuate player performance. Paradoxically, it rewards inefficient players. However, in spite of such criticisms, official competitions continue to use the system.

(4) Adequacy of the MVP selection system of the ACB Spanish League
The main competition in Spain is called ACB League. The ACB League gives value to the performance of players by giving out different recognitions and awards. It is the competition using a more heterogeneous system to designate MVP. Usually, leagues designate the most valuable player of the whole season. In the ACB League, together with the MVP of the season, there are recognitions given to the "Player of the month" and "Player of the week".

The MVP of the season is selected by a panel of experts together with the votes of fans. The MVP of the month is selected only by a panel of experts. And finally, the MVP of each week is selected only by statistical judgement (the player who has lead the "Ranking" score that week). Consequently, there is a mixture of methods in designating such recognitions.

 © 2011 University of Alicante

(5) Relationship between player evaluation metrics and market value

Berri, Brook and Schmidt (2007) show how player salary is related to scoring in the NBA. But scoring is only a simple indicator of player performance. The logical assumption made is that players with high statistical ratings will have increased media exposure, and hence market value. In addition, the use of advanced statistics by specialized analysts may favour some underrated players, highly proficient players but with lesser marketing surround them. This is the case for example, of Daryl Morey, General Manager of the Houston Rockets since 2006. Morey is a convinced defender of contracting players based on their advanced (non obvious, i.e., beyond box score) stats. A proof of his ability to find underrated players is the performance of Houston in the last two seasons (2009 and 2010). Rockets finished both regular seasons with a positive winning record, competing in the harder conference of the NBA, with a testimonial participation of the two "franchise players" of the team, Yao Ming and Tracy McGrady, because of injuries. Any other team would have collapsed given the same situation. Probably, a similar situation would have collapsed any other team[6].

But what is the perception of stakeholders regarding this issue in Euroleague and Spanish ACB League? Recall that in Europe, advanced statistical analysis is not so popular than in the USA. Although this is a growing area (webs like www.inthegame.org or www.draftexpress.com are starting to provide advanced stats), the situation is still far from the USA. Note that using one or other technique to valuate player is not a trivial issue to rank players. As Martínez (2010c) shows, even highly correlated systems may yield important divergences when players are ordered according to their value. Therefore, using one specific system could help or harm a player from the viewpoint of media exposure and public recognition.

(6) Knowledge and preferences for some of the most widely used evaluation systems, and use of advanced stats.

Do stakeholders know the advantages of using advanced stats? Do they know sufficiently the multiple options available to valuate players? This is an important concern in the dissemination and use of player evaluation metrics. We have evaluated knowledge and preferences of stakeholders regarding a sample of methods to evaluate player. The aim was to describe the current awareness, understanding, liking and use of some of the most known systems.

If stakeholders do not know the possibilities of statistics they could be reluctant to use it. Do they think that "Ranking" outperforms other more advanced stats such as, for example, "Player Efficiency Rating" (Hollinger, 2005) or "Adjusted Plus Minus" (Winston, 2009). The analysis of responses of stakeholders may help to know if they are near or far from this quantitative revolution.

(7) The best methods to assess the value of a player.

Finally, which is the best technique or set of techniques to assess player value? Currently, there is a plethora of statistical and mathematical methods used by the creators of the systems. These are just a sample: linear regression, data envelopment analysis, optimization via ridge regression, probabilistic estimation using a finite population approach, cluster analysis, financial mathematics using the Black-Scholes approach, etc. Obviously, there are other proposals that are not such mathematically demanding, relying only on theory, experience or simple deduction. Each method has advantages and shortcomings, but it would be interesting to explore the opinions of a specific group of experts about the feasibility of using certain type of techniques.

[6] *A simple example about the philosophy of Morey can be found in Ballard (2009) or Lewis (2009).*

In sum, stakeholder opinions regarding these key issues may serve to evaluate the current situation of basketball analytics regarding player evaluation metrics, helping to define what has to be measured and how has to be measured. With a special focus on Euroleague and Spanish ACB League, this research will help to calibrate the perception of stakeholders regarding the evaluation of players, and whether their evaluation could imply marketing effects.

MATERIAL AND METHODS

Data collection and measures

The first step was to identify stakeholders. The selection criteria for the study participants included: (1) an intensive relationship with the problem to be evaluated, or (2) a proven background in academic research. Note that not all of the stakeholders would answer to all of the proposed research questions. Therefore, we identified the following groups of interests: players (active or inactive professional players), coaches (with experience in professional leagues), agents (with experience with professional players), editors and journalists (specialized in basketball), analysts (working for a professional team or for a mass media), chairs (highly proficient academics of the following disciplines: marketing, business, economy, sport and management), researchers (highly skilled researchers with published papers on basketball analytics in a Thompson-Scientific journal), bloggers (fans with an actualized blog about basketball), and fans (highly active basketball lovers).

The chair group needs an additional explanation. Chairs did not have to necessary know the world of basketball, but were selected because in marketing, economy, business or management, the concept of value is a recurrent topic (the value of a customer, the value of a brand, the value of a stock, the value of a relationship, etc.). Therefore they could help to solve some of the questions proposed, adding an interdisciplinary view, what it was also desirable.

We designed a procedure for data collection based on the construction of a large data base with potential sample units. As contact would be realized by email, it was necessary to obtain a valid email address for each potential participant. Although some of the research questions were specific to Euroleague and Spanish basketball, we did not limit the study to European participants. In terms of the Spanish participants, the cultural link with the authors and the nature of some questions of the study were advantageous. Therefore, we focused the search of a portion of all the potential sample units only in the Spanish ambit, such as fans and chairs. In addition, we also focused the main search of bloggers in the Spanish context, although we also explored a great amount of blogs around the world.

A database was created with more than 1700 potential sample units, which matched with the criteria established to participate in the study. We needed several weeks to explore a plethora of basketball sites in the internet, in order to collect the email addresses. We also used the site www.eurobasket.com to register the email addresses of all the teams (only teams in the first division of each country) appearing in its database (the most complete team database available), and also the emails of international agents. In order to contact coaches and players we also used the ACB League site (www.liga-acb.es) and the Spanish Basketball Federation (www.feb.es). The Spanish Federation supported the study by disseminating the aim of this research to the coaches of the Federation.

In the search for chairs, we navigated for the websites of all the Spanish universities in order to get a valid contact form. For the fans, we used the most important basketball forum in Spain, the official forum of the Spanish ACB League. We requested permission from the ACB League to use the shared email addresses

 © 2011 University of Alicante

of registered users (more than 40,000 users were registered at that time). We selected the first 875 users (they had posted at least 500 messages to the forum), but only 278 emails addresses were publicly shared.

In the selection of researchers, we searched for basketball papers published by Thompson-Scientific Journals. We reviewed some of the most outstanding journals which included papers about basketball: Journal of Sport Management, European Sport Management Quarterly, Journal of Sports Economics, European Journal of Sports Science, Economic Inquiry, European Journal of Operational Research, Journal of Sports Science, among others. The aim was to identify proficient researchers with knowledge in basketball analytics. We also wrote a post in the APBRmetrics site in order to call the attention of APBRmembers. The aim was to invite participants of the APBR site to accomplish the questionnaire. The APBR community comprises the most talented professional quantitative analysts, because most of the analysts who currently work or have worked for NBA teams have participated in this forum[7].

Editors and journalists specialized in basketball were selected from magazines, e-magazines and corporate mass-media sites around the world, with special attention to Spanish stakeholders. Players were the most difficult group to access, and we used personal websites, personal blogs and mass-media relationship to contact with them.

It must be noted that there is no clear distinction between groups of interests. Recall that, for example, many coaches have been former players, many journalists are also bloggers or specialized analysts, agents can also be scouts and former players, members of the APBRcommunity can be editors or professional analysts working for NBA teams and all of them are basketball fans. For example one of the participants is currently a coach, but was a former player in the national team of his country, winning a medal in Summer Olympics. This is an important issue to consider in analysing their opinions.

A resume of the sample used is showed in Table 1. The categories of stakeholders are only a way to describe the sample, because participants can pertain to several categories at time. However, the criteria we chose to classify participants in those groups were in line with the main professional activity of these individuals. For example, if a journalist is also a blogger, we categorize him as journalist, if a former professional player is also a coach or a blogger, we categorize him as player, if a blogger collaborates with a mass media but he is not a journalist then we categorize him as blogger, etc.

The procedure to obtain the stakeholders opinions was the following. A personalized email was sent to each potential participant asking for his help to collaborate in a research about basketball analytics. There was a guarantee of anonymity; we could not make public the names or email addresses of participants. We ensured the participants of their copies of the results of the research upon publication. An attached text file was sent with the questions to provide additional information. For example, for the chairs group, an additional text was included in order to familiarize chairs with the problem being studied before they answer the questionnaire. A post about the study was published in the blog of one of the authors, in order to motivate the participation of site visitors. Participants had several weeks to accomplish the questionnaire and send it to the authors via e-mail. To facilitate stakeholder participation, the questionnaire was prepared in two versions: English and Spanish. Data collection started in May of 2009 and finished in April of 2010. Mail-delivery error was about 8% of the total sample. Thus, a significant amount of emails did not reach the target recipients. The error rate may be attributed to errors in the email database of Eurobasket.com and

[7]*As we have commented previously, the revolution of the quantitative analysis in basketball is a phenomenon that comes from United States, and there are much more researchers and analysts in that country with respect to the rest of the world.*

errors in emails provided for the own site of many teams. In addition, anti-spam filters were also a threat to our data collection.

As Table 1 shows, 182 questionnaires were received; so the response ratio was about 10%. The asymmetry of response ratio among groups is remarkable. The worst response ratio was found among coaches because the vast majority of emails to them was no directly sent to their personal e-mail addresses but to the corporate email addresses of those teams (using the www.eurobasket.com database). The prevalence of Spanish participants in all the categories was also evident, except for the researchers and analysts. In the category of analysts we only considered APBRmetrics members. We did not directly contact the members, but we wrote a post to the APBRmetrics forum encouraging the participation of members. We did not consider all the members as potential sample units (there are more than 1200 members in the APBR community).

Although we cannot reveal the identities of participants, we may highlight interesting features of some of them. For example, (1) among players, there were 4 components of the Spanish National Team in the last fifteen years; (2) among coaches, there were 2 head coaches of the Spanish National Team, and 2 head coaches who won supranational competitions coaching ACB teams; and (3) among APBRmembers there were 4 participants that have worked for NBA teams as analysts. However, we must also stress that opinions were not weighted in function of the characteristic of participants, but they were aggregated, i.e. each response contributed in the same way to the global analysis.

Participants answered questions without space limitations. They could use all the time and space that they needed to reply. We calculated that the time necessary to accomplish the questionnaire ranged from 5 to 20 minutes. The number of words written by participants was about 56000.

Regarding measures, a small introductory text preceded the questions being formulated. In addition, for evaluating the knowledge and preferences for some of the most widely used evaluation systems, a list of some indexes were provided. Participants had to valuate using a 0-10 scale their preferences for each one of the indexes, but only if they know enough of each one. Therefore, if they did not know enough of an index they were not to evaluate it. In order to control for potential response bias (for example, participants assessing an index that they did not know enough), we included some consistency questions. For example, the system called “NBA Efficiency” (used by the NBA) is computed in the same way that the index called “System of evaluation of the French League” (used by the French basketball league). Consequently, if both systems were valuated, they had to be the same score, in order to consider a valid response. The complete list of systems was: “Ranking”, “SEDENA”, “System of evaluation of French League”, “Tendex” “NBA Efficiency”, “Player Efficiency Rating”, “Net Plus Minus”, “Adjusted Plus Minus”, “Manley Credits”, “IBM Award”, “Mays Magic Metric”, “Points created”, “Steele Value”, “Hoopstat Grade”, “Diamond Rating”, “Points responsible”, “Total Performance Rating”, “Data Envelopment Analysis”, “Berri’s Individual Wins”. These systems are explained in Martínez (2010a), and they are only a small set of all systems that are currently available for quantitative analysis. Within this set of indices, there are the most used in official competitions around the world (Martínez, 2010b).

 © 2011 University of Alicante

Table 1. *Description of the sample*

	Players	Nº	Coaches	Nº	Agents	Nº	Editors and Journalists	Nº	Bloggers	Nº	Researchers	Nº	Analysts	Nº	Chairs	Nº	Fans	Nº	Total	Nº	%
Sample	19		700		127		112		51		43		8		426		278		1764		
Replies	6		20		4		37		29		12		8		47		19		182		
Reponse ratio (%)	31.58		2.86		3.15		33.04		56.86		27.91		100.00		11.03		6.83		10.32		
Nationality	Spain	6	Spain	14	USA	2	Spain	27	Spain	27	Spain	4	USA	7	Spain	47	Spain	19	Spain	144	79.12
			Ireland	1	Italy	1	Argentina	2	France	1	South Korea	1	Italy	1					USA	19	10.44
			Italy	1	Greece	1	Bosnia	1	Italy	1	USA	6							Italy	5	2.75
			Lietuva	1			USA	4			Greece	1							Greece	3	1.65
			New Zealand	1			Ghana	1											Argentina	2	1.10
			Portugal	1			Greece	1											Bosnia	1	0.55
			Switzerland	1			Italy	1											France	1	0.55
																			Ghana	1	0.55
																			Ireland	1	0.55
																			Lietuva	1	0.55
																			New Zealand	1	0.55
																			Portugal	1	0.55
																			South Korea	1	0.55
																			Switzerland	1	0.55
(1) The definition of the "value of a player"	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes			182	
(2) The role of qualitative evaluations	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes			182	
(3) The adequacy of the "Ranking" system.	Yes		Yes		Yes		Yes		Yes		Yes		Yes		No		Yes			135	
(4) Adequacy of the MVP selection system of the ACB Spanish League	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes			182	
(5) Relationship between player evaluation metrics and market value	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes			182	
(6) Knowledge and preferences for some of the most widely used evaluation systems, and use of advanced stats.	Yes		Yes		Yes		Yes		Yes		Yes		Yes		No		Yes			135	
(7) The best methods to assess the value of a player.	No		No		No		No		No		No		No		Yes		No			47	

Data analysis

We followed the principles of descriptive content analysis (Neuendorf, 2002; Riffe, Lacy and Fico, 2008) to analyse stakeholder's opinions. Once the research questions were proposed, the next step was to establish categories to code opinions. We followed a mixture of deductive and inductive method (Neuendorf, 2002), proposing a priori categories but keeping flexible to add new categories if they emerged from the process of codification. Table 2 shows the a priori categories defined for all research questions. Therefore, we prioritized participants voice (Thompson, 1997; Kelly and Tian, 2004), combining the deductive approach with the flexibility of the inductive method, in an effort to incorporate some advantages of interpretative research to the formalism of quantitative content analysis.

Regarding the definition of the value of a player, we proposed four categories based on an extensive literature review (e.g. Berri, Brooks and Schmidt, 2006; Martínez, 2010a; Oliver, 2004; Winston, 2009). The first category referred to the contribution of the player to team winning, and the second the offensive and defensive performance on the court (including effective decisions, and intangibles such as leadership, intensity, etc.). Obviously both categories are very related but the first only included responses that not named intangible factors. In addition, the second category did not include the "winning" aspect. The third category referred to actions on the court relative to their team role. This included responses mentioning that value had to be related with the role of the player in the team (position of the court, expectations of coach, etc.). Finally, the fourth category referred to normalize actions on the court considering contextual factors (league average players, clutch time, minutes and possessions, etc.). Certainly, participants could include in their definition all the categories. This means that categories overlapped (an undesired situation under the view of Riffe et al., 2008), but it was a necessary requirement for this type of question. Responses were classified in any possible combination of the four categories, i.e. 12 possible outcomes, so that we could categorize responses from small definitions (e.g. 10 words), to large definitions (e.g. 50 words). Large definitions were more prone to be categorized as a combination of categories.

In terms of the role of qualitative evaluations, we asked about the superiority or inferiority of qualitative evaluations regarding quantitative evaluations. Therefore, we proposed 5 categories: (1) complement of quantitative evaluations; (2) better than quantitative evaluations, or preponderance of qualitative vs. quantitative; (3) worse than quantitative evaluations, or preponderance of quantitative vs. qualitative; (4) do not rely on qualitative evaluations made by journalists; (5) rely on qualitative evaluations made by journalists. Again, responses were classified in any possible combination between the three first categories (they are mutually exclusive) and the latter two categories (which are also mutually exclusive), i.e. 11 possible outcomes.

Regarding the adequacy of the "Ranking" system, we proposed two categories: (1) the system is adequate, i.e. stakeholders agree with the current use of this system; (2) the system needs to be improved, because it has several limitations. We only considered in the first category those opinions which no expressed an explicit criticism.

Regarding the adequacy of the MVP selection system in the ACB League, we created four disparate categories; (1) the stakeholders agree with the current system; (2) they criticize the use of the "Ranking"; (3) they criticize the use of a panel of experts; (4) they do not agree about using fan vote. Again combinations of the second, third and fourth categories were possible, because criticisms could be possible in a combination of categories.

 © 2011 University of Alicante

Regarding the relationship between player evaluation metrics and player market value, we simply categorized responses in a Yes/No format. We relied on the explanations of the respondents to know the nuances of this simple categorization.

Regarding knowledge, preference and use of advanced stats, participants had to valuate using a 0-10 scale their preferences for each one of the indexes. In addition, we proposed a question to Spanish participants (except players, researchers and chairs) and international agents, regarding the use of advanced stats (other than Ranking) in their diary work (for coaches, journalists and bloggers) and as fans. The aim was to know not only the preference or familiarity with other indexes but if they effectively use them. The categories were the following: (1) Use Ranking; (2) Do not use Ranking; (3) Use other advanced stats.

Finally, we inquired chairs regarding the more proper method or family of methods to assess the value of a player. We wanted to consider a well grounded opinion regarding the potential applications of methods such as for example: data envelopment analysis, neural nets, supervised learning, etc. We did not inquire analysts or researchers about that because they are familiar with the advanced methods and they use some of them in their specific proposals. However, as it was mentioned earlier, we wanted an interdisciplinary view provided by researchers with the highest academic degree and profile. Categories proposed were the following: (1) Only quantitative methods; (2) Only qualitative methods: (3) Mixture of quantitative and qualitative methods[8]. In the first category were included responses defending linear weights and statistics and data mining techniques (regression, neural nets, k-nearest neighbours, discriminant analysis, etc.). The second category included responses that only mentioned the qualitative evaluations by a panel of experts. In the third category were included qualitative evaluations based on quantitative analysis, global evaluations based on considering quantitative and qualitative information, and techniques which combine both approaches, such as data envelopment analysis (see Cooper, Ruiz and Sirvent, 2009) or multicriteria analysis (see Saaty, 2008). Table 2 summarizes all the categories.

We present the results of data analysis in the following section. These results are shown for the complete actual sample, and for different groups of stakeholders. We decided to cluster some groups of stakeholders as a function of some common criteria. Therefore, we grouped players, coaches and agents (C1) because they are directly implied in the valuation of players. Players and coaches are part of the game, and agents are responsible to bargain contracts for players. Editors, journalists and bloggers are in the second cluster (C2). They share many characteristics: they analyze the game, write columns, discuss viewpoints, interview players, etc. Some could argue that bloggers are not "professionals", so they cannot be compared with journalists, but the sampled bloggers are responsible for their blogs, and perform very similar tasks to journalists (although they are not paid for their work and are not subject to some type of editorial review). Researchers and analysts are the last group (C3). They have specific and proficient skills in basketball research. They are well-grounded in statistical methods, and publish studies (in academic and not

[8]*Regarding the quantitative/qualitative distinction, it is apparent (and several researchers and chairs pointed out this issue) that there is no purely quantitative analysis, because qualitative information is always needed. For example, in the construction of linear weights ratings such as Efficiency or Ranking, qualitatively is assumed that the weights are the same for all variables. Even in the econometrics model, such as Berri (1999), there is a qualitative assumption based on what is the dependent variable to build the model (wins) and which are the predictors (see the view of Pearl, 2000, regarding the role of qualitative assumptions to these types of models). However, we maintain the labels to distinguish among disparate methods that are very different in its nature.*

academic sources) using quantitative analysis applied to basketball. All of them may be considered as researchers, but we draw the distinction between academic participants and non-academic participants[9].

Table 2. *Research questions and categories.*

Research questions	**Categories**
(1) The definition of the "value of a player"	1. Contribution to team winning 2. Offensive plus defensive performance (effective decisions) including intangibles (leadership, psychology, etc.) 3. Actions on the court relative to their team role (position, expectation of the coach) 4. Actions on the court relative to context (average player, plays, rivals, etc.) *5. Box-score* *6. Fan satisfaction and team revenues* *7. Only intangible factors*
(2) The role of qualitative evaluations	1. Complement of quantitative evaluations 2. Better than quantitative evaluations or preponderance of qualitative vs. quantitative 3. Worse than quantitative evaluations or preponderance of quantitative vs. qualitative 4. Do not rely on qualitative evaluations made by journalists 5. Rely on qualitative evaluations made by journalists *6. Should be done by recognized experts*
(3) The adequacy of the "Ranking" system.	1. Adequate 2. Needs to improve
(4) Adequacy of the MVP selection system of the ACB Spanish League	1. Adequate 2. Criticisms about using the Ranking system 3. Criticism about using such panel of experts 4. Criticism about using fan votes *5. Marketing focused* *6. Same criteria for all designations*
(5) Relationship between player evaluation metrics and market value	1. Yes 2. No
(6) Knowledge and preferences for some of the most widely used evaluation systems, and use of advanced stats.	1.Use Ranking 2. Do not use Ranking 3. Use other advanced stats
(7) The best methods to assess the value of a player.	1. Only quantitative methods 2. Only qualitative methods 3. Mixture of quantitative and qualitative methods

Note: Italics represent no a priori categories, i.e. emergent categories which were added in the process of codification

[9]*Actually this distinction is not pervasive, because some academic researchers are very close to some professional teams, and some non-academic researchers are starting to publish in academic sources.*

 © 2011 University of Alicante

As the process of assigning responses to categories was done independently by the two authors, we present intercoder reliability for each research question, using Percent Agreement, Kappa coefficient[10] (see Riffe et al., 2008), Krippendorf Alpha[11] (Krippendorf, 2004) and Spearman correlation[12] (Schmidt and Hunter, 1996)[13]. All of the reliability coefficients computed was above 0.8. Coefficients above this threshold are generally considered as sound, i.e. the process of codification was reliable (Krippendorff, 2004). However, it is clear that measurement errors existed, and this influenced comparisons between cells in the contingency tables created.

In terms of statistical data analysis, we focused our interest in comparing responses between some specific groups, using the two-group comparison approach recommended by Rothman (1990), Savitz and Olsan (1998), Mayo and Cox (2006), and Rothman, Greenland and Lash (2008), instead of the multiple group approach associated with multiway contingency tables. Therefore, we used Phi coefficient as a measure of effect size (Grissom and Kim, 2005). This coefficient ranges from -1 to 1, with zero indicating no association, and it is equal to Cramer's V in 2x2 or 2xn tables. It is specially recommended in naturalistic sampling, where the researchers have chosen only the total sample size, and not the row or column sample sizes, so that any variation between the two column total or between the two row totals is natural rather than based on the researcher's arbitrary choices of sample sizes (Grissom and Kim, 2005). Obviously, this is the case of this research where group's sample sizes were not a priori determined.

As we stated, measurement error influenced effect sizes computed. This is an important issue in social science (Hunter and Schmidt, 2004). Several authors have recommended some guidelines to correct effect sizes by measurement error (e.g. Schmidt and Hunter, 1996; Ree and Carreta, 2006), but these guidelines do not cover effect sizes coefficient for contingency tables. The form of how this bias acts is highly complicated, because it depends on how the agreement is distributed among cells. We may analyse this effect using a heuristic method based on simulation. As Kappa is equal to Phi in 2x2 tables when agreement is above that would be expected by chance, then we may simulate the effect of measurement error under different circumstances. For example, the percentage of agreement (PA) in a 2x2 table (where cells are A1B1, A1B2, A2B1, A2B2) is PA= (A1B1 + A2B2) / total sample. In addition, the ratio between diagonal cells (RDC) is RDC = A1B1/A2B2. Using these simple indexes, we ran several simulations that are showed in Figure 1, using a total sample of 182 units, as in the case of our empirical study. The assumption underlying this simulation is that responses in non-diagonal cells are mistaken, due to the codification process.

[10]This coefficient has been criticized by some authors, such as Krippendorff (2004). This is one of the reasons we used several coefficients for measuring intercoding agreement.

[11]We implemented Hayes and Krippendorff (2007) procedure to compute Krippendorff Aplha, providing 95% confidence interval for estimates. The macro written for those authors were run in SPSS for those computations.

[12]Although Spearman correlation should be used only where there are ordered categories instead of nominal categories (Neuendorf, 2002), we used this coefficient as a proxy of non-parametric association between coders, which is also a proxy of the shared variance between coders. Spearman correlation, Kappa coefficient and Pearson correlation between coders are very similar for all variables involved in our study.

[13]An example of the interminable debates about some statistical methods is the use of chi-square to fit structural equation models. This simple statistic has been a matter of subject in several books (Hayduk, 1996) and special issues of journals such as Structural Equation Modeling: A Multidisciplinary Journal, or Personal and Individual Differences.

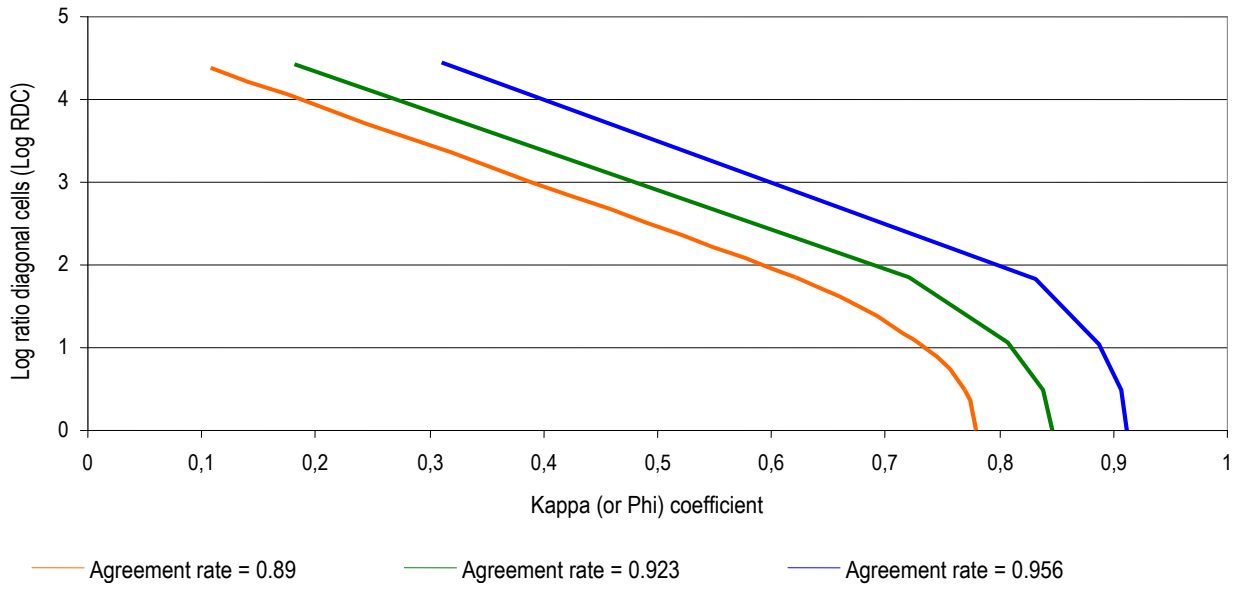

Figure 1. *Simulation on the effect of measurement error on reliability.*

As can be seen, Kappa (or Phi) increases with PA with a similar RDC. When RDC increases, diagonal cells are more dissimilar and this fact decreases Kappa (or Phi). Note that RDC is presented in a logarithmic scale to facilitate graphical representation, i.e. using the original scale, the effect would be much more prominent. Consequently, it was very difficult to analyse how measurement error influenced effect size (Phi) in our study. Although we obtained high levels of intercoder reliability (Spearman correlation, Krippendorf Alpha, Kappa and PA were above 0.8), we decided to be conservative in the estimation of reliability. Our reasoning is consistent with the view of Schmidt and Hunter (1996). These authors explain how high intercoder agreement could be artificially inflated because of several sources of error (random, specific and transient). These errors could be controlled by administering the questionnaire on different occasions, but this was unpractical for our research. Recall that measurement error underestimate coefficient of association such as Phi, that it is in fact a Pearson correlation for dichotomous data. Therefore we reported results using the uncorrected effect sizes and the corrected effect sizes. In the first case, as intercoder reliability was very high in our study, these effect sizes would reflect an optimistic scenario, where random, specific and transient error would be negligible. In the second case, effect sizes would reflect a conservative scenario, where those sources of errors would be present. We decided to fix reliability to 0.8 in this latter case. This analysis of scenarios of reliability has been used for other authors as a form of sensitivity analysis (Martínez-Costa, Choi, Martínez and Martínez-Lorente, 2009).

 © 2011 University of Alicante

RESULTS

(1) The definition of the "value of a player"

As Table 3 shows, the vast majority of responses (almost 94%) matched with the *a priori* proposed categories. However, three new categories emerged from the process of codification: (5) Box-score, i.e. player value is what box-score statistics reflect, similar to the Ranking system; (6) Fan satisfaction, i.e. player value should be defined in terms of fulfilment of expectations of fans; (7) Only intangible factors, i.e. player value is only linked with factors such as intensity or leadership. These three new categories have a very marginal representation, so we may limit the value definition to the four *a priori* categories we proposed.

Table 3. *Distributions of responses for (1) The definition of the "value of a player"*

	Total sample	**C1**	**C2**	**C3**	**C4**	**C5**
1. Contribution to team winning	64 (35.95%)	10 (33.33%)	21 (32.81%)	8 (42.10%)	6 (31.57%)	19 (41.30%)
2. Offensive plus defensive performance (effective decisions) including intangibles (leadership, psychology, etc.)	116 (65.16%)	25 (83.33%)	48 (75.00%)	8 (42.10%)	11 (57.89%)	24 (52.17%)
3. Actions on the court relative to their team role (position, expectation of the coach)	48 (26.96%)	6 (20.00%)	14 (21.87%)	6 (31.57%)	4 (21.05%)	18 (39.13%)
4. Actions on the court relative to context (average player, plays, rivals, etc.)	43 (24.15%)	10 (33.33%)	19 (29.68%)	4 (21.05%)	1 (5.26%)	9 (19.56%)
5. Box-score	8 (4.49%)	0 (0%)	3 (4.68%)	0 (0%)	0 (0%)	5 (11.1%)
6. Fan satisfaction	4 (2.24%)	0 (0%)	1 (1.56%)	1 (5.26%)	0 (0%)	2 (4.34%)
7. Only intangible factors	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (5.26%)	0 (0%)
Valid responses	178 (97.8%)	30 (100%)	64 (96.96%)	19 (95.00%)	19 (100%)	46 (97.87%)

Percentage of agreement: 0.972. Kappa: 0.966. Spearman correlation: 0.99. Krippendorff Alpha: 0.966 (0.963; 0.993). *Note: Percentages are computed considering only valid responses. As responses could be coded in multiple categories, the sum of percentages exceeds 100%. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C3= Researchers and analysts. C4= Fans. C5= Chairs.*

About two thirds of the participants (65.16%) specifically linked value to offensive and defensive performance on the court, including intangible factors, such as leadership, psychology, etc. Linking value to wins, or normalizing by other factors are not such prevalent responses. Among clusters of stakeholders, there were no important differences in the response pattern. Table 4 shows effect sizes for the four main categories. However, it is important to note that responses of cluster 1 (players, coaches and agents) and cluster 2 (editor, journalists and bloggers) were almost identical (Phi=0.040). In addition, researcher and analysts (cluster 3) and chairs (cluster 5) responded in a very similar fashion (Phi=0.070). Therefore, some differences (although with a low effect size) were encountered among clusters 1 and 2, and clusters 3 and 5. It seems that analysts and academics (researchers and chairs) were more prone to linked value with wins.

Table 4. *Effect sizes for the four main categories of (1) The definition of the "value of a player".*

Effect sizes	(C1)	(C2)	(C3)	(C4)	(C5)
(C1)		0.040	0.279	0.261	0.279
(C2)	0.032		0.206	0.190	0.242
(C3)	0.223	0.165		0.293	0.070
(C4)	0.209	0.152	0.235		0.208
(C5)	0.223	0.194	0.056	0.166	

Note: Uncorrected effect sizes are in the lower-triangle and corrected effect sizes are in the upper triangle. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C3= Researchers and analysts. C4= Fans. C5= Chairs.

(2) The role of qualitative evaluations

As Table 5 shows, qualitative evaluations should be a complement of quantitative evaluations for almost half of the participants (45.50%). It is noteworthy to indicate that there was no consensus regarding the preponderance of qualitative vs. quantitative evaluations or vice versa, because there is the same percentage of opinions defending one vs. the other. In addition, about one of each four participants expressed non-reliance on the judgements of journalists who qualitatively valuate players. Not one of the participants specifically stated that they were comfortable with the criteria of journalists. This is one of the reasons that one category emerged from the process of codification. Therefore, some participants (7.86%) specifically recognized that those qualitative evaluations should be done by true experts, and not some journalists who have little background in basketball.

 © 2011 University of Alicante

Table 5. *Distributions of response for (2) The role of qualitative evaluations.*

	Total sample	C1	C2	C3	C4	C5
1. Complement of quantitative evaluations	81 (45.50%)	13 (43.33%)	34 (53.12%)	14 (70.00%)	7 (36.84%)	23 (50.00%)
2. Better than quantitative evaluations or preponderance of qualitative vs. quantitative	36 (20.22%)	3 (10.00%)	15 (23.43%)	5 (25.00%)	4 (21.05%)	8 (17.39%)
3. Worse than quantitative evaluations or preponderance of quantitative vs. qualitative	35 (19.66%)	7 (23.33%)	10 (15.62%)	1 (5.00%)	6 (31.57%)	11 (23.91%)
4. Do not rely on qualitative evaluations made by journalists	47 (26.40%)	15 (50.00%)	18 (28.12%)	2 (10.00%)	7 (36.84%)	5 (10.86%)
5. Rely on qualitative evaluations made by journalists	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
6. Should be done by recognized experts	14 (7.86%)	3 (10.00%)	2 (3.12%)	1 (5.00%)	1 (5.26%)	7 (15.21%)
Valid responses	178 (97.8%)	30 (100%)	64 (96.96%)	20 (100%)	19 (100%)	46 (97.87%)

Percentage of agreement: 0.938. Kappa: 0.915. Spearman correlation: 0.96. Krippendorff Alpha: 0.915 (0.865; 0.957). *Note: Percentages are computed considering only valid responses. As responses could be coded in multiple categories, the sum of percentages exceeds 100%. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C3= Researchers and analysts. C4= Fans. C5= Chairs.*

Regarding the response pattern by clusters (Table 6), what is remarkable is the disparity of thinking of cluster 1 and the remaining clusters, especially with respect to analysts and academics (clusters 3 and 5). Players, coaches and agents, were more reluctant to rely on qualitative evaluations instead of quantitative ones. The first reason for this thinking is the distrust of the criteria of journalists.

Table 6. Effect sizes for the four main categories of (2) The role of qualitative evaluations.

Effect sizes	C1	C2	C3	C4	C5
C1		0.276	0.548	0.217	0.431
C2	0.221		0.259	0.276	0.242
C3	0.438	0.207		0.544	0.300
C4	0.173	0.172	0.435		0.322
C5	0.345	0.194	0.240	0.258	

Note: Uncorrected effect sizes are in the lower-triangle and corrected effect sizes are in the upper triangle. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C3= Researchers and analysts. C4= Fans. C5= Chairs.

Comparing Table 5 and 6, we may see that effect sizes are higher in the second case. This indicates that differences in the opinions among clusters were more noteworthy for the second research questions.

(3) The adequacy of the "Ranking" system.
Majority of participants (71.64%) acknowledged that the "Ranking" system needs improvement (Table 7). There was a remarkable similarity among the responses of clusters 1 and 2, and the responses of clusters 3 and 4 (Table 8). This means that players, coaches, agents, editors, journalists and bloggers were much more comfortable with "Ranking" system than analysts, researchers and fans. Clusters 3 and 4 were much more aware about the limitations of this rating system than other clusters.

Table 7. *Distributions of response for (3) The adequacy of the "Ranking" system.*

	Total sample	C1	C2	C3	C4
1. Adequate	38 (28.35%)	11 (37.93%)	21 (31.81%)	3 (15.00%)	3 (15.78%)
2. Needs to improve	96 (71.64%)	18 (62.06%)	45 (68.18%)	17 (85.00%)	16 (84.22%)
Valid responses	134 (98.52%)	29 (96.66%)	66 (100%)	20 (100%)	19 (100%)

Percentage of agreement: 0.932. Kappa: 0.904. Spearman correlation: 0.909. Krippendorff Alpha: 0.904 (0.809; 0.980). *Note: Percentages are computed considering only valid responses. As responses could be coded in multiple categories, the sum of percentages exceeds 100%. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C3= Researchers and analysts. C4= Fans.*

© 2011 University of Alicante

Table 8. *Effect sizes for the four main categories of (3) The adequacy of the "Ranking" system.*

Effect sizes	C1	C2	C3	C4
C1		0.074	0.311	0.297
C2	0.069		0.198	0.185
C3	0.249	0.158		0.013
C4	0.238	0.148	0.011	

Note: Uncorrected effect sizes are in the lower-triangle and corrected effect sizes are in the upper triangle. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C3= Researchers and analysts. C4= Fans.

(4) Adequacy of the MVP selection system of the ACB Spanish League

The MVP selection system of the ACB Spanish League needs to improve. As Table 9 shows, only 40.11% of the participants considered this system as adequate. The main criticism to this system is the use of fan votes. However, about 10% of the participants recognized that this is desirable from a marketing viewpoint, so this is one of the categories that emerged from the analysis. Other participants (7.18%) specifically stated the necessity of using the same criteria for all designations. And this was the other category that emerged. It is also noticeable that researchers and analysts (cluster 3) did not criticize the use of the panel of experts to design the MVP. These stakeholders expressed their opinions in a more different fashion than the other clusters, as Table 10 shows.

Table 9. *Distributions of response for (4) Adequacy of the MVP selection system of the ACB Spanish League.*

	Total sample	C1	C2	C3	C4	C5
1. Adequate	67 (40.11%)	13 (52.00%)	20 (31.74%)	6 (33.33%)	9 (50.00%)	19 (44.18%)
2. Criticisms about using the Ranking system	38 (22.75%)	7 (28.00%)	19 (30.15%)	2 (11.11%)	3 (16.66%)	7 (16.27%)
3. Criticism about using such panel of experts	28 (16.76%)	3 (12.00%)	13 (20.63%)	0 (0%)	5 (27.77%)	7 (16.27%)
4. Criticism about using fan votes	65 (38.92%)	6 (24.00%)	27 (42.85%)	9 (50.00%)	8 (45.11%)	15 (34.88%)
5. Marketing focused	17 (10.17%)	3 (12.00%)	5 (7.93%)	4 (22.22%)	1 (5.55%)	4 (9.30%)
6. Same criteria for all designations	12 (7.18%)	0 (0%)	2 (3.17%)	0 (0%)	4 (22.22%)	6 (13.95%)
Valid responses	167 (91.75%)	25 (83.33%)	63 (95.45%)	18 (90.00%)	18 (94.73%)	43 (91.48%)

Percentage of agreement: 0.958. Kappa: 0.950. Spearman correlation: 0.991. Krippendorff Alpha: 0.950 (0.914; 0.985). Note: Percentages are computed considering only valid responses. As responses could be coded in multiple categories, the sum of percentages exceeds 100%. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C3= Researchers and analysts. C4= Fans. C5= Chairs.

Table 10. *Effect sizes for the four main categories of (4) Adequacy of the MVP selection system of the ACB Spanish League.*

Effect sizes	C1	C2	C3	C4	C5
C1		0.253	0.609	0.286	0.202
C2	0.203		0.302	0.184	0.205
C3	0.487	0.241		0.409	0.318
C4	0.229	0.147	0.327		0.096
C5	0.162	0.164	0.255	0.077	

Note: Uncorrected effect sizes are in the lower-triangle and corrected effect sizes are in the upper triangle. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C3= Researchers and analysts. C4= Fans. C5= Chairs.

(5) Relationship between player evaluation metrics and market value

As Table 11 clearly shows, stakeholders believed that there is a relationship between player evaluation metrics and player value (83.70%). This means that quantitative metrics are important for bargaining contracts and for the media projection of players. Regarding effect sizes, Table 12 shows that player, coaches and agents (cluster 1) have a more conservative view comparing with researchers, analysts and chairs. These two latter clusters (3 and 5) are almost absolutely convinced regarding the influence of the quantitative ratings on the market value of a player.

Table 11. *Distributions of response for (5) Relationship between player evaluation metrics and market value.*

	Total sample	C1	C2	C3	C4	C5
1. Yes	149 (83.70%)	21 (72.41%)	52 (80.00%)	17 (85.00%)	15 (78.94%)	44 (97.77%)
2. No	29 (16.30%)	8 (27.59%)	13 (20.00%)	3 (15.00%)	4 (21.06%)	1 (2.22%)
Valid responses	178 (97.80%)	29 (96.66%)	65 (98.48%)	20 (100%)	19 (100%)	45 (95.74%)

Percentage of agreement: 0.994. Kappa: 0.980. Spearman correlation: 0.980. Krippendorff Alpha: 0.979 (0.939; 1.000). *Note: Percentages are computed considering only valid responses. As responses could be coded in multiple categories, the sum of percentages exceeds 100%. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C3= Researchers and analysts. C4= Fans. C5= Chairs.*

© 2011 University of Alicante

Table 12. *Effect sizes for the four main categories of (5) Relationship between player evaluation metrics and market value.*

Effect sizes	C1	C2	C3	C4	C5
C1		0.105	0.185	0.092	0.473
C2	0.084		0.067	0.013	0.327
C3	0.148	0.054		0.098	0.306
C4	0.074	0.011	0.079		0.400
C5	0.379	0.262	0.245	0.321	

Note: Uncorrected effect sizes are in the lower-triangle and corrected effect sizes are in the upper triangle. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C3= Researchers and analysts. C4= Fans. C5= Chairs.

(6) Knowledge and preferences for some of the most widely used evaluation systems, and use of advanced stats.

As we showed in Table 7, only 28.35% of stakeholders considered the Ranking system as adequate. However, as Table 13 shows, the majority of them (73.19%) use this system in their diary work (cluster 1 and 2) or in their discussion or own analysis of basketball (fans). One of the reasons of this behaviour may be that these stakeholders admitted that they did not use other advanced stats (only 13.40% handle advanced metrics). However, coaches and agents are more familiar with these advanced stats and use them in a more significant fashion than editors, journalists, bloggers and fans (Table 14). It seems clear that stakeholders do not use other advanced stats because they do not know enough these advanced stats (Table 15). Only Efficiency is known to the majority of stakeholders (67.86%), although this percentage is not very high considering that Efficiency is a simple variation of Ranking, i.e., they yield almost identical results. PER, Net Plus Minus and Adjusted Plus Minus are known approximately by the half of participants, but the remaining systems had very low representation.

Table 13. *Distributions of response for (6) Knowledge and preferences for some of the most widely used evaluation systems, and use of advanced stats.*

	Total sample	C1	C2	C4
1.Use Ranking	71 (73.19%)	18 (75.00%)	38 (71.69%)	15 (78.94%)
2. Do not use Ranking	19 (19.58%)	2 (8.33%)	13 (24.52%)	4 (21.06%)
3. Use other advanced stats	13 (13.40%)	8 (33.33%)	4 (7.54%)	0 (0%)
Valid responses	97 (97.00%)	24 (100%)	53 (100%)	19 (100%)

Percentage of agreement: 0.957. Kappa: 0.911. Spearman correlation: 0.915. Krippendorff Alpha: 0.911 (0.822; 0.977). Note: Percentages are computed considering only valid responses. As responses could be coded in multiple categories, the sum of percentages exceeds 100%. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C4= Fans.

Table 14. *Effect sizes for the four main categories of 6) Knowledge and preferences for some of the most widely used evaluation systems, and use of advanced stats.*

Effect sizes	C1	C2	C4
C1		0.416	0.833
C2	0.329		0.184
C4	0.667	0.147	

Note: Uncorrected effect sizes are in the lower-triangle and corrected effect sizes are in the upper triangle. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C4= Fans.

Table 15. *Knowledge and preferences for some of the most widely used evaluation systems.*

	Total sample	Knowledge	C1	Knowledge	C2	Knowledge	C3	Knowledge	C4	Knowledge
1. Ranking	6.97	98.21%	7.11	96.43%	7.21	92.98%	5.15	65.00%	7.41	94.44%
2. SEDENA	7.26	23.21%	7.33	21.43%	7.19	28.07%		0%	7.50	22.22%
3. System of evaluation of French League	7.71	6.25%	8.00	17.86%	7.00	3.51%		0%		0%
4. Tendex	6.32	30.36%	7.00	28.57%	6.23	22.81%	5.83	60.00%	8.00	5.56%
5. NBA Efficiency	6.96	67.86%	7.50	57.14%	7.11	66.67%	5.17	60.00%	7.70	55.56%
6. Player Efficiency Rating	7.25	55.36%	7.20	53.57%	7.30	49.12%	7.00	65.00%	7.67	33.33%
7. Net Plus Minus	6.52	46.43%	7.00	42.86%	6.35	40.35%	6.25	60.00%	6.80	27.78%
8. Adjusted Plus Minus	6.99	42.86%	7.00	35.71%	6.73	35.09%	7.54	65.00%	6.60	27.78%
9. Manley Credits	5.78	12.50%	7.50	7.14%	5.86	12.28%	4.75	20.00%	6.00	5.56%
10. IBM Award	5.88	20.54%	7.00	17.86%	5.67	15.79%	5.29	35.00%	6.00	11.11%
11. Hoopstat Grade	6.23	11.61%	7.00	7.14%	6.00	8.77%	6.40	25.00%	5.00	5.56%
12. Diamond Rating	4.14	6.25%		0%	3.67	5.26%	4.00	15.00%	6.00	5.56%
13. Steele Value	6.17	15.18%		0%	6.18	19.30%	6.00	25.00%	7.00	5.56%
14. Belloti Points Created	6.00	21.43%	6.75	14.29%	5.55	19.30%	6.00	35.00%	7.00	11.11%
15. Clearbout Quality Points	6.14	6.25%	8.00	3.57%	5.67	5.26%	6.00	10.00%	6.00	5.56%
16. Mays Magic Metric	6.00	6.25%	8.00	3.47%	6.00	3.51%	5.67	15.00%	5.00	5.56%
17. Total Performance Rating	7.00	8.93%	8.00	7.14%	7.00	8.77%	6.33	15.00%		0%
18. Berri Individual Wins	6.00	24.11%	6.89	10.71%	6.13	15.79%	6.88	65.00%	4.50	11.11%
19. Points responsible	6.81	16.96%	8.50	7.14%	6.69	14.04%	6.38	40.00%	8.00	5.56%
20. Data Envelopment Analysis	7.00	12.50%		0%	7.33	10.53%	6.86	35.00%	6.00	5.56%
Valid responses	110	81.48%	27	90.00%	54	81.81%	14	70.00%	17	89.47%

Note: Percentages are computed considering only valid responses. As responses could be coded in multiple categories, the sum of percentages exceeds 100%. 12 responses were dropped because the failed to pass the consistency test. C1= Players, coaches and agents. C2= Editor, journalists and bloggers. C3= Researchers and analysts. C4= Fans.

© 2011 University of Alicante

Researchers and analysts know much more the advanced stats than the other stakeholders. This fact would be predictable when comparing with fans, but it is also evident when comparing with the other stakeholders. Therefore, coaches, agents and specialized mass media workers have a less background in the field of player rating systems than researchers and analysts.

Regarding the preferences for the different rating systems, the low sample sizes made data analysis difficult because standard errors were high for many of the estimations. However, there was some interesting information found in the responses of participants. For example, there is no perfect rating system, because the mean value of the most cited systems is about seven points in a scale that ranged from 0 to 10. Globally, PER is the preferred system (7.25), followed by Adjusted Plus Minus (6.99), Ranking (6.97), Efficiency (6.96) and Net Plus Minus (6.52). Anyway, the magnitude of the differences among them, i.e., effect size is very low, almost negligible, so we could argue that they are valued in a very similar fashion. Note that this is an interesting result because Adjusted Plus Minus and Net Plus Minus are completely different systems from the rest.

Nevertheless, if we analyse clusters responses we find that researchers and analysts are much more critical with Ranking (5.15) than the other stakeholders. As expected, a similar valuation is done for Efficiency (5.17). Adjusted Plus Minus and PER are valuated significantly better (7.44 and 7.00, respectively). In addition, this cluster is the only one that knows the Berri's Individual Wins. However, they did not valuate Berri's work in a significantly better way (6.88) than Adjusted Plus Minus or PER.

(7) The best methods to assess the value of a player.
Finally, chairs expressed their opinions regarding the most proper method or set of methods to assess the value of a player. As Table 16 shows, quantitative methods are not enough to correctly establish a rating system for valuating players. A mixture of quantitative and qualitative methods is preferred (56.09%). However, it is noticeable to stress that very opposite opinions were found regarding the use of quantitative vs. qualitative methods. Several chairs stressed their disconformities with the use of human judgement to valuate player, whilst other chairs defended the use of qualitative methods. This is consistent with results showed in Table 9, where the same percentage of chairs (16.27%) criticized the use of a quantitative system as Ranking and the use of a panel of experts.

Table 16. *Distributions of response for (7) The best methods to assess the value of a player.*

	Total sample	**Chairs**
1. Only quantitative methods	19.51%	19.51%
2. Only qualitative methods	24.39%	24.39%
3. Mixture of quantitative and qualitative methods	56.09%	56.09%
Valid responses	41 (87.23%)	41 (87.23%)

Percentage of agreement: 0.952. Kappa: 0.918. Spearman correlation: 0.925. Krippendorff Alpha: 0.919 (0.798; 1.000). Note: Percentages are computed considering only valid responses.

DISCUSSION

In this research we examined the opinions of basketball stakeholders regarding several questions of special interests to valuate players. Players, coaches, agents, journalists, editors, bloggers, researchers, analysts, fans and chairs participated in this macro-research. After analysing their opinions using the content analysis methodology, we discuss the main results of this research.

Players should be mainly valued in terms of their offensive and defensive contribution on the court, considering intangible elements such as leadership, intensity, intimidation, personality, etc. Taking this principle into account, the definition of value could be improved by adding the normalization factor, i.e. to normalize such evaluation by context factors, such as the comparison of a league average player, expectations of the coach, position, possessions, minutes played, strength of schedule, clutch plays, etc.). About one of each three stakeholders specifically linked value to wins. When stakeholders spoke about offensive and defensive contribution to team, they also meant performance that is linked to team success, i.e. the win factor. Therefore, it would be also desirable to prize players who help their teams win.

There are rating systems such as Wins Produced (Berri, 1999; Berri and Schimidt, 2010), Win Shares (Kubatko, 2009) and others "wins ratings" (see Martinez and Martinez 2010a) that match those requirements of the stakeholders. However, there is a problem with these metrics. If we centre our analysis on Wins Produced (the only rating system published in a peer-review journal), we see that box-score statistics explain about 94% of the variance of team wins (after adjusting by team specific factors). But our research has showed that stakeholders think that intangible factors are also very important. Therefore, if some combination of box-score statistics can predict wins with only a 6% of error, then intangibles should not count up to 6% of the variations in wins. And this is not what stakeholders think. However, the accuracy of the model of Wins Produced is very laudable. Note that the sum of player's individual wins predicts team wins with a very low error rate (Berri and Schmidt, 2010). Consequently, do stakeholders overestimate intangibles? Is the model specification of Wins Produced flawed? Maybe the solution is to separate production from factors facilitating production. Wins Produced explain almost perfectly wins (and wins are based on score difference between teams), but intangibles such as leadership, commitment, etc., are factors facilitating scoring of the own team and preventing scoring of the opponent team. Using a psychometric analogy, player performance would be a latent variable which is partially reflected by an observable indicator: productivity derived from box-score statistics.

As Berri's work is not very much appreciated by researchers and analysts with respect to the other metrics. In addition, other authors, such as Winston (2009) consider that 80% of the game is outside the box-score, although he provides no discussion of how he came up with this 80% figure. This is one of the ideas behind his Adjusted Plus Minus, and other versions of this rating system (see Martínez and Martinez 2010a). But again, the criticism to this family of Plus Minus metrics have been clearly detailed by Berri and Bradbury (2010) and Berry and Schmidt (2010). Our research has also showed that this family of metrics were not valued in a more significant way than others. We do not think that Berri's model was misspecified, but intangibles are, in the majority of cases, factors that facilitate offensive and defensive performance. Consequently, for stakeholders, box-score statistics represent tangible performance, and are insufficient to correctly valuate players.

There is no index that can measure intangibles in the form that stakeholders think. Again Plus Minus indexes have several limitations and they lack consistency and stability (Berri and Bradbury, 2010). A solution would be to consider the subjective criteria of experts, in order to complement quantitative

 © 2011 University of Alicante

evaluations. Almost half of stakeholders agree with this statement, but the problem is that around 20% of participants think that qualitative evaluations are not very much reliable. This percentage is similar to the percentage of stakeholders that do not rely on the use of an expert panel to select the MVP and to the percentage of chairs that think that only quantitative methods are adequate to study the value of players. The use of other experts distinct from journalists would help to minimize this disagreement. Several stakeholders have expressed the need of using "true" experts, such as specialized analysts, coaches or players who would make more fair qualitative evaluations. However it is also true that some journalists are well recognized specialists, but stakeholders think (especially players, coaches, agents, analyst and researchers) that the vast majority of journalists provide biased qualitative evaluations.

Another solution would be to create a system that combines several ratings, in order to overcome the limitations of each individual rating. Several indexes such as S.C.O.R.E (www.goldenstateworriers.com/2009/12/warriors-scoreboard-12-21-09.html, Composite Score (Nichols, 2009a) or Roland Rating (www.82games.com/0809/ROLRTG8.HTM) combine several indexes (including Plus Minus stats), but the problem is how to combine the systems that form the composite one, because those criteria are somewhat arbitraries. In addition, other indexes such as Z-Points (www.binahoop.com) are intelligent systems to capture the importance of each action in the game. This system matches with the requirements of several participants which demand a way to weight game actions in terms of clutch minutes and tied score.

Ranking system is a deficient system. Almost four of each five stakeholders agree with this. As stated before, criticisms of this linear system are well-known among analysts and researchers, but we have found that the majority of the remaining stakeholders think in a similar way. Furthermore, the percentage of responses criticizing Ranking would probably increase if stakeholders would know other rating systems better. This is one of the most interesting results we have found. Knowledge of other systems to valuate player is generally poor among non-analysts and researchers stakeholders. Note that the use of other advanced stats by stakeholders is very limited. Therefore, researchers and analysts are the stakeholders who have better knowledge of the different alternatives to valuate players, so these are the individuals with more background to understand the limitations of Ranking and the ways this system could be improved. The problem again is that there is no agreement to decide which better system it is. Researchers and analysts are comfortable with Hollinger's PER (Hollinger, 2005), but criticism on this system are also well known (Berri, 2010, Berri and Bradburi, 2010; Winston, 2009). The correction made by Treutlein (2009) to PER (the Alternate Player Efficiency Rating, APER), overcome some of the limitations of PER, so maybe it would be an attractive rating system to apply for PER defenders.

Although the Ranking system is a deficient rating system, surprisingly around a 40% of stakeholders agree with its use to choose the weekly MVP. However, the motive for this paradoxical result is related with practical reasons. Several stakeholders would prefer a mixture of quantitative and qualitative methods, but they stated that because of time restrictions and immediate designation of weekly MVP, a quantitative rating system would be desirable. However, there are a plethora of easily understandable and applicable rating systems based on linear weights, which could be used instead of Ranking. Martínez (2010a;c) reviews them. The main criticism to the MVP designation system is the use of fan votes. However, about 10% of the participants recognized that there are marketing reasons to count fan's opinions. It is true that fan's participation is marginal in the designation of MVP around the different leagues of the world. As Martínez (2010b) shows, only in the leagues of France, Germany, Philippines and Somalia (together with the ACB League and Euroleague) do fans participate in the MVP designation. We think that marketing

efforts of such competitions are laudable, but stakeholders think that it would be fairer to give two MVP awards: the expert designation and the fan designation.

Therefore, what is the best form of choosing the MVP of a league? It seems clear that a mixture of quantitative and qualitative methods would be desirable. In analyzing player value, stakeholders prefer this combination of methods. Note that although several competitions use the quantitative rating systems to select MVP, such as the leagues of Spain, Balkan, Baltic, Bulgaria, Slovak Republic, Eurochallenge, Eurocup, Euroleague, Luxemburg, Portugal and Serbia, the majority of these designations refer to the weekly or monthly MVP. The vast majority of competitions around the world use a panel of experts to choose the league MVP (Martínez, 2010b). The prevalence of only this qualitative designation is evident, and after analysing stakeholder's responses, the best form of making such qualitative decision would be the use of players or coaches vote, instead of mass media participation. In leagues such as BJ-League in Japan, VTB League in East Europe, D-League in USA and others (see Martínez 2010b), journalists do not participate in the panel of experts to designate the MVP of the season. Coaches, players or both, make such decision. It is important to note that chairs preferred the combination of both quantitative and qualitative methods to assess the value of a player. Basically, they commented on two options: (1) to make a qualitative decision based on quantitative ratings, i.e., adding value to the numeric rating; or (2) to use Data Envelopment Analysis (DEA) to choose the most efficient player based on some qualitative rules. The first option would be easier to apply, and experts would consider quantitative information and then other intangible factors to make the decision. The second option would be necessarily applied by people trained in linear programming. In addition, there are several chairs that specifically named DEA as an unreliable technique. The disagreement among chairs regarding mathematical techniques was obvious. For example, some of them defended neural nets as a proper technique to handle several statistical information from players, but others argued that neural nets are generally over parameterized and they are very opaque. Consequently, one of the problems to implement a sophisticated mathematical method to analyse player value is the conflicting views regarding the advantages and shortcomings of the different methods. This is a common problem in science, and this is one of the factors that restrict the advance of several disciplines. Multicriteria analysis (Saaty, 2008) was one of the innovative proposals that never have been applied to this field. However, the most promised proposal was related with only quantitative analysis, taking box-score statistics as compositional data (Aitchison, 1982; Pawlowsky-Glahn, and Egozcue, 2001), where the vector representing player statistics would have a relative nature, based on the global vector of the team stats. Further research should deepen into this innovative proposal.

Finally, there is consensus among stakeholders regarding the marketing benefits obtained by players who are best valued. However, potential marketing losses are suffered by players who are not recognized using the official system of the league (such as Ranking in ACB or Euroleague) but who would otherwise be recognized using another type of advanced metric. Several participants stated that this is more important in the NBA than in Europe competitions. A remarkable number of players, coaches and analysts think that rating systems are not such important to revalorize or to sign players, and to bargain contracts, because they think that decision makers (team general managers, coaches, etc.) do not base such decisions in statistical outcomes. In addition, there is a trend in the responses of some stakeholders to undervalue statistics, especially some players, coaches, journalists and fans. This is an important fact, because it seems that some basketball people do not rely on the contribution of statistics to the game of basketball. Contrary to the statistical social revolution emerged in the last years, they are reluctant to incorporate statistics to make decisions. We interpret that some of them feel threaten by the irruption of people who believe and use statistics for decision making. However, these sceptic individuals do not know enough the

 © 2011 University of Alicante

advantages of quantitative analysis in basketball, so, in the majority of cases, they do not have the background enough to criticize rating systems and the contribution of statistical analysis to basketball.

CONCLUSION

As Oliver (2004) and other gurus of basketball statistics acknowledge, there is no Holy Grail of stats. There is no single number that could say all about the value of a player. However, using advanced rating systems is a way to approach this chimera. The better the rating system, the better the approximation to the perfect number, and the better objectivity gained in the valuation of players.

This research has shown the view of stakeholders regarding player evaluation. Therefore, player value can be defined as the offensive and defensive contribution of the player on the court, considering intangible elements such as leadership, intensity, intimidation, personality, etc. In addition, this contribution has to be linked to winning the game, the main objective of professional sports. In order to achieve comparison among players, it is necessary a normalization process, weighting factors such as minutes, possessions, average player of the league, etc.

Box-score statistics can be used to measure player contribution to wins with a high level of accuracy. However, these statistics do not reflect intangibles. As intangibles are a critical factor to stakeholders in order to valuate a player, player ratings systems are biased towards players with a tangible contribution. There is no player rating system that correctly considers intangibles, because Plus Minus indexes provide inconsistent outcomes. Therefore, human qualitative judgement would improve the assessment of a player, but it would be necessary that these evaluations be made by true experts, and this excludes the majority of journalists and fans.

We encourage competitions to use other rating methods instead of Ranking. Although there is no perfect quantitative method, there is ample evidence about the shortcomings of Ranking compared with other systems, such as, for example, Wins Produced. One of the reasons of this reluctance to change could be the poor knowledge of stakeholders (excluding researchers and analysts) regarding the plethora of alternatives existing in the basketball analytic world. Therefore, in order to be fairer in the quantitative player valuation, competitions should incorporate other rating systems. One of the consequences of this hypothetical change would be the influence of the market value of the players, because some statistical classifications and MVP designation would also change.

ACKNOWLEDGEMENTS

We are totally in debt with each one of the 182 participants that voluntarily responded the questionnaire.

REFERENCES

1. AITCHISON J. The statistical analysis of compositional data (with discussion). *Journal of the Royal Statistical Society*, *Series B (Statistical Methodology)*. 1982; 44(2):139-177. [Abstract] [Back to text]
2. ARIELY D. *Predictably Irrational: The Hidden Forces That Shape Our Decisions.* HarperCollins; 2008. [Abstract] [Back to text]
3. BALLARD C. *The art of a beautiful game*. Simon & Shuster: New York; 2009. [Abstract] [Back to text]
4. BARFIELD O. *Saving the appearances; Study in idolatry* (2nd edition). Middletown, CT: Wesleyan University Press; 1988. [Abstract] [Back to text]

5. BERGER PL, LUCKMANN T. *The social construction of reality: A treatise in the sociology of knowledge*.Garden City, NY: Doubleday; 1966. [Abstract] [Back to text]
6. BERRI DJ. Who is 'most valuable'? Measuring the player's production of wins in the National Basketball Association. *Managerial and Decision Economics*. 1999; 20:411-427. doi:10.1002/1099-1468(199912)20:8 [Back to text]
7. BERRI DJ. Measuring performance in the National Basketball Association. *Working paper;* 2010. [Back to text]
8. BERRI DJ, BROOK SL, SCHMIDT MB. Does One Simply Need to Score to Score? *International Journal of Sport Finance.* 2007; 2(4):190-205. [Abstract] [Full Text] [Back to text]
9. BERRI DJ, BRADBURY JC. Working in the land of metricians. *Journal of Sports Economics*. 2010; 11(1):29-47. doi:10.1177/1527002509354891 [Back to text]
10. BERRI DJ, SCHMIDT MB. *Stumbling on wins: Two economists expose the pitfalls on the road to victory in professional sports*. FT Press; 2010. [Abstract] [Back to text]
11. BERRI DJ, SCHMIDT MB, BROOK SL. *The wages of wins: Taking measure of the many myths in modern sport*. Palo Alto,CA: Stanford University Press; 2006. [Abstract] [Back to text]
12. COLDITZ GA, ATWOOD KA, EMMONS K, MONSON RR, WILLETT WC, TRICHOPOULOS D, HUNTER DJ. Harvard Cancer Risk Index. *Cancer Causes Control 11.* 2000; 477-488. [Abstract] [Full Text] [Back to text]
13. COOPER WW, RUIZ JL, SIRVENT I. Selecting non-zero weights to evaluate effectiveness of basketball players with DEA. *European Journal of Operational Research*. 2009; 195(2):563-574. doi:10.1016/j.ejor.2008.02.012 [Back to text]
14. DOOLITTLE B, PELTON K. *Pro Basketball Prospectus 2009-10*. Prospectus Entertainment Ventures LLC; 2009. [Abstract] [Back to text]
15. DOYLE P. *Marketing management and strategy.* Fourth edition. Financial Times/ Prentice Hall; 2006. [Abstract] [Back to text]
16. ESTELLER-MORÉ A, ERES-GARCÍA M. A note on consistent players' valuation. *Journal of Sports Economics*. 2002; 3(4):354-360. doi:10.1177/152700202237500 [Back to text]
17. FRANSELLA F, BELL R, BANNISTER R. A *manual for repertory grid technique.* Second Edition. Wiley; 2004. [Abstract] [Back to text]
18. GIGERENZER G, TODD PM, & THE ABC RESEARCH GROUP. *Simple Heuristics That Make. Us Smart*. New York: Oxford University Press; 1999. [Back to text]
19. GRISSOM RJ, & KIM JJ. *Effect sizes for research. A broad practical approach*. Mahwah, NJ: LEA; 2005. [Abstract] [Back to text]
20. HAYDUK LA. *LISREL Issues, Debates and Strategies*. Ed. Johns Hopkins University Press: Baltimore, MD; 1996. [Abstract] [Back to text]
21. HAYES AF, KRIPPENDORFF K. Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures.* 2007; 1:77-89. [Abstract] [Full Text] [Back to text]
22. HOLLINGER J. *Pro Basketball Forecast*. Washington, D.C.: Potomac, Inc; 2005. [Abstract] [Back to text]
23. HUNTER JE, SCHMIDT FL. *Methods of Meta-Analysis: Correcting Error and Bias in Research. Findings*. 2nd Edition. Newbury Park: Sage Publications; 2004. [Abstract] [Back to text]
24. JOHN DJ, LOKEN B, KIM KH, MONGA AB. Brand Concept Maps: a methodology for identifying brand association networks. *Journal of Marketing Research.* 2006; 43(4):549-563. [Abstract] [Full Text] [Back to text]
25. KELLY SW, TIAN K. Fanatical consumption. An investigation of the behaviour of sports fans through textual data. In Kahle Riley (Ed), *Sports marketing and the psychology of marketing communication.* 2004; 27-65. [Abstract] [Back to text]

 © 2011 University of Alicante

26. KRIPPENDORFF K. *Content analysis: an introduction to its methodology*. 2nd ed. Sage publication. Thousand Oask, California; 2004. [Abstract] [Back to text]
27. Kubatko J. Calculating Win Shares. Retrieved from http://www.basketballreference.com/about/ws.html; 2009. [Full Text] [Back to text]
28. KUKSOV D, VILLAS-BOAS JM. When more alternatives lead to less choice. *Marketing Science.* 2010; 29(3):507-524. [Full Text] [Back to text]
29. LEVITT SD, DUBNER SJ. *Freakonomics: A rogue economist explores the hidden side of everything*. New York: William Morrow; 2005. [Abstract] [Back to text]
30. LEWIN D, ROSENBAUM DT. The pot calling the kettle black. Are NBA statistical models more irrational than "irrational" decision-makers? *New England Symposium on Statistics in Sports*. Harvard University Science Center; 2007. [Full Text] [Back to text]
31. LEWIS, MM. *Moneyball: The art of winning an unfair game*. W.W. Norton & Company Inc; 2003. [Abstract] [Back to text]
32. LEWIS MM. *The no-stats all stars*. Downloaded from http://www.nytimes.com/2009/02/15/magazine/15Battier-t.html; 2009. [Full Text] [Back to text]
33. MARTÍNEZ JA. Una revisión de los sistemas de valoración de jugadores de baloncesto (I). Descripción de los métodos existentes. *Revista Internacional de Derecho y Gestión del Deporte.* 2010a; 10:37-77. [Abstract] [Full Text] [Back to text]
34. MARTÍNEZ JA. Una revisión de los sistemas de valoración de jugadores de baloncesto (II). Competiciones oficiales y ligas de fantasía. *Revista Internacional de Derecho y Gestión del Deporte.* 2010b; 11:48-68. [Abstract] [Full Text] [Back to text]
35. MARTÍNEZ JA. Una revisión de los sistemas de valoración de jugadores de baloncesto (III). Discusión general. *Revista Internacional de Derecho y Gestión del Deporte.* 2010c; 12:44-79. [Abstract] [Full Text] [Back to text]
36. MARTÍNEZ JA, MARTÍNEZ L. Un método probabilístico para las clasificaciones estadísticas de jugadores en baloncesto. *Revista Internacional de Ciencias del Deporte.* 2010d; 18(6):13-36. doi:10.5232/ricyde2010.01802 [Back to text]
37. MARTÍNEZ M, CHOI T, MARTÍNEZ JA, MARTÍNEZ AR. ISO 9000/1994, ISO 9001/2000 and TQM: The performance debate revisited. *Journal of Operations Management.* 2009; 27(6):495-511. doi:10.1016/j.jom.2009.04.002 [Back to text]
38. MAYO DG, COX DR. Frequentists statistics as a theory of inductive inference. In Rojo, J. (ed.), *2nd Lehmann Symposium- Optimally. IMS Lecture Notes-Monographs Series*. 2006; 1-28. doi:10.1214/074921706000000400 [Back to text]
39. NEUENDORF KA. *The content analysis guidebook*. Sage Publications. Thousand Oaks, London; 2002. [Abstract] [Back to text]
40. NICHOLS J. Explanation of Composite Score. Retrieved from http://basketballstatistics.com/aboutcs.html; 2009. [Full Text] [Back to text]
41. O'DONOGHUE PG, WILLIAMS J. An evaluation of human and computer-based predictions of the 2003 rugby union world cup. *International Journal of Computer Science in Sport.* 2004; 3(1):5-22. [Abstract] [Back to text]
42. OLIVER D. *Basketball on paper. Rules and tools for performance analysis.* Washington, D. C.: Brassey's, INC; 2004. [Abstract] [Back to text]
43. PAWLOWSKY-GLAHN V, EGOZCUE JJ. Geometric approach to statistical analysis on the simplex. *Stochastic Environmental Research and Risk Assessment (SERRA).* 2001; 15(5):384-398. doi:10.1007/s004770100077 [Back to text]
44. PEARL J. *Causality: Models of Reasoning and Inference*. Cambridge University Press; 2000. [Abstract] [Back to text]

45. PIETTE J, ANAND S, ZHANG K. Scoring and shooting abilities of NBA players. *Journal of Quantitative Analysis in Sports.* 2010; 6(1) Article 1. doi:10.2202/1559-0410.1194 [Back to text]
46. REE MJ, CARRETA TH. The role of measurement error in familiar statistics. *Organizational Research Methods.* 2006; 9(1):99-112. doi:10.1177/1094428105283192 [Back to text]
47. RIFFE D, LACY S, FICO F. *Analyzing media messages. Using quantitative content analysis in research*. Taylor & Francis Group. New York; 2008. [Abstract] [Back to text]
48. ROTHMAN KJ. No adjustments are needed for multiple comparisons. *Epidemiology.* 1990; 1(1):43-46. [Abstract] [Back to text]
49. ROTHMAN KJ, GREENLAND S, LASH TL. *Modern Epidemiology.* Third edition. Philadelphia: Luppincott, Williams & Wilkins; 2008. [Abstract] [Back to text]
50. SAATY TL. Relative Measurement and Its Generalization in Decision Making Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors The Analytic Hierarchy/Network Process. *Review of the Royal Spanish Academy of Sciences*. 2008; 102(2):251-318. [Abstract] [Full Text] [Back to text]
51. SAVITZ DA, OLSHAN AF. Decribing data requires o adjustment for multiple comparisons: a reply form Savitz and Olshan. *American Journal of Epidemiology.* 1998; 147:813-814. [Abstract] [Back to text]
52. SCHMIDT FL, HUNTER JE. Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods.* 1996; 1(2):199-223. [Back to text]
53. SKINNER B. The Price of Anarchy in Basketball. *Journal of Quantitative Analysis in Sports*. 2010; 6(1) Article 3. [Abstract] [Full Text] [Back to text]
54. THOMPSON CJ. Eureka! and other tests of signifcance: A new look at evaluating interpretive research, in *Advances in Consumer Research* Volume 17, eds. Marvin E. Goldberg and Gerald Gorn and Richard W. Pollay, Provo, UT : Association for Consumer Research. 1990; (25-30). [Full Text] [Back to text]
55. THOMPSON CJ. Interpreting Consumers: A Hermeneutical Framework for Deriving Marketing Insights from the Texts of Consumers Consumption Stories, *Journal of Marketing Research.* 1997; 34(Nov.):438-55. [Abstract] [Back to text]
56. TREUTLEIN J. *APER: Player Efficiency Ranking adjusted for Assisted Field Goals.* Retrieved from http://www.hoopdata.com/recent.aspx?aid=39; 2009. [Back to text]
57. WINSTON WL. *Mathletics.* New Jersey: Princeton University Press; 2009. [Abstract] [Back to text]
58. ZALTMAN G. *How Customers Think: Essential Insights into the Mind of the Markets*. Boston: Harvard Business School Press; 2003. [Abstract] [Back to text]

 © 2011 University of Alicante