THE BASKETBALL EVALUATION SYSTEM

(A computerized factor weighted model with measures of validity)

Kenneth L. Swalgin

Pennsylvania State University, USA

Original scientific paper
UDC: 796.323:621.317
Received: February 20, 1998
Accepted: May 16, 1998

Abstract

In the United States, coaches at all levels of play have devised player evaluation systems. As early as 1941, Foster "Phog" Allen, the coach who replaced James Naismith at the University of Kansas, developed a value weighted factor rating system to determine the effectiveness of offensive and defensive performance (Elbell and Allen, 1941). This system and all others to date have posed a number of problems. They are either too cumbersome to administer, are based on subjective and/or outdated criteria, and/or do not distinguish between positions of play (Swalgin, 1987). The author wishes to propose a more scientific approach to player evaluation.

The aim of this research was to enhance the Basketball Evaluation System (BES), a computerized performance evaluation model which grades player performance in relationship to "position of play" and "time played" under game conditions. Weighting factors for an established set of performance criteria were developed from a survey of expert coaches. The weighting factors were then incorporated into the original model in an attempt to strengthen the validity of scores produced for overall performance. Scores for the unweighted and weighted models were then correlated with a set of criterion scores established from another group of expert coaches. The results indicate that both models correlated highly with the coaches' criterion scores. For the unweighted model, r = .757 and for the weighted model, r = .798. The difference between the correlations was not statistically significant. The addition of factor weighting did however add to the face validity of the model, giving coaches a quantitative tool to measure individual performance in relationship to "position of play" and "time played".

Key words: basketball, performance, evaluation, computers, models

Zusammenfassung:

BASKETBALLBEWERTUNGSSYSTEM (Computerisiertes Modell der Belastungsfaktoren mit Validitätsmessung)

In den Vereinigten Staaten haben die Trainer an allen Wettkampfsstufen die Systeme zur Spielerbewertung entwickelt. Schon in 1941 entwickelte Foster "Phog" Allen, der Trainer, der James Naismith an der Kansas Universität vertauscht hat, ein System der wertbelasteten Faktoren, mit dem Ziel, die Effizienz beim Angriffs- und Verteidigungsspiel festzustellen (Elbell und Allen, 1941). Dieses System, wie auch alle anderen bis heute haben viele Probleme gestellt. Sie sind entweder ungünstig für die Anwendung, oder basieren auf subjektiven oder veralteten Kriterien und/oder unterscheiden nicht zwischen den Spielpositionen (Swalgin, 1987). Der Author wollte einen wissenschaftlicheren Zugang zur Spielerbewertung vorschlagen.

Das Ziel der Untersuchung war, das Basketballbewertungssystem (BES), ein computerisiertes Leistungsbewertungsmodell, das die Spielerleistung im Bezug auf die "Spielposition" und die "Spielzeit" bewertet, zu vollenden. Die Belastungsfaktoren für eine bestimmte Reihe der Leistungskriterien wurden mittels einer Ankete unter den Experttrainern entwickelt. Die Belastungsfaktoren wurden dann in das originelle Modell inkorporiert, im Versuch, die Validität der Ergebnisse für die Gesamtleistung zu verbessern. Dann wurden die Ergebnisse für das belastete sowie für das unbelastete Modell mit den Ergebnissen für die Kriterienreihe einer anderen Trainergruppe verglichen. Die Ergebnisse zeigen, daß beide Modelle mit den Ergebnissen der Trainerskriterien hoch korreliert haben. Für das unbelastete Modell r = 0,757 und für das belastete r = 0,798. Der Unterschied zwischen den Korrelationen ist statistisch nicht bedeutend. Die Faktorenbelastung hat doch dem Nominalwert des Modells beigetragen, was den Trainern ein quantitatives Meßmittel zur individuellen Spielerleistung im Bezug auf die "Spielposition" und die "Spielzeit" anbietet.

Schlüsselwörter: Basketball, Leistung, Bewertung, Computer, Modell

Introduction

The process to effectively improve team and individual performance often centers upon the coach's ability to observe, measure and analyze performance. This process (Swalgin, 1994) commonly includes: (a) defining, categorizing, and prioritizing

performance factors inherent to the sport, (b) determining the degree of success or failure of the performance, and then (c) effectively implementing an intervention method to improve the quality of play. This evaluation process is not only essential as a means to evaluate players in the short term, but also to establish expectations for consistent performance in the long term. Mike Krzyzewski, the Duke University basketball coach (personal communication, May 4, 1989), stated that "evaluation of game performance is essential for coaches looking for consistent performance."

To give coaches a tool to effectively measure the performance elements that lead to successful play, The Basketball Evaluation System (BES) was devised (Swalgin, 1992)[1]. The BES statistical performance evaluation model is based on "Normal Probability Theory" and grades player performance in relationship to "position of play" and "minutes played" under game conditions. Burson (1973) states that, "it is a certainty that the real measure of a player or team's achievement is their performance under game situations. Ability to perform skills outside the game situation is not a guarantee that these same skills will be performed with equal effectiveness in the game itself."

The BES model incorporates three measurement constructs that can be applied to most team sport structures: (a) a common set of performance criteria specific to the sport, (b) a norm based context to measure the criteria, and (c) an accurate, functional measurement system inherent to the structure of the sport. According to Deming (1975), to develop an effective system of evaluation, four requirements must be met:

- that the system have a meaningful operational measure of success or failure
- that there be a satisfactory design to test or evaluate the data
- that the results lead to action different from the action that would be taken on the basis of the original data
- that there be a group of people authorized to take action regarding the results.

The BES model grades nine game related performance factors, referred to as Scaled Performance Scores (SPS) and generates an overall grade called the Graded Performance Score (GPS). The model standardizes raw game scores and grades each factor on a scale ranging from 0 to 4, similar to the academic grading structure used in the United States. The computerized model produces grades after each game, and on a season-to-date basis. The graded performance factors include: Field Goal percentage (FG%), Three Point Field Goal percentage (3FG%), Free Throw percentage (FT%), Rebounds (REB), Personal Fouls (PF), Assists (AST), Turnovers (TO), Blocked Shots (BLK), and Steals (STL).

The validity of the scores produced by the model were established for individual performance factors and overall performance in an earlier study (Swalgin, 1993). Correlation coefficients between BES scores and a set of criterion scores established from a group of expert coaches (n = 14) were calculated. For individual performance factors, the scores range from r = .481 for BLK to r = .902 for STL. For overall performance, r = .837. In this study, to further test the variability between the scores produced for overall performance, a correlation matrix was calculated. Here, each coach's overall rating was correlated with the BES overall ratings, and with all other coaches. From the correlation matrix, the average correlation was determined between the BES overall ratings and the coaches' scores. For BES, the average correlation was 0.70, with a range from (0.43 - 0.93). The average correlation between all coaches' were also determined. Here, the average correlation equaled 0.60, with a range from (0.19 - 0.96). These findings indicate that the Basketball Evaluation System shows less variance than the coache's

[1]Support for this study was partially funded by a grant from the National Association of Basketball Coaches (NABC). To date, three versions of the software have been developed: college men, college women, and high school boys. The software is IBM compatible and a DOS application. For more information on the Basketball Evaluation System, feel free to write, fax, or e-mail the author. Dr. Ken Swalgin, Box 222, Pennsylvania State University, York, PA 17403 USA. Fax (717) 771-8404, e-mail KXS1@PSU.EDU

ratings when combining factors to produce an overall rating.

To establish the measurement context for the current model, national performance norms representing 70% of all players by "position of play" in men's Division I (one) college basketball were developed in 1994. Norms were established in relationship to position of play and time played. The positions include; Point Guard (PG), Off Guard (OG), Small Forward (SF), Power Forward (PF), Center © and the "swing" positions, PG-OG, OG-SF, SF-PF, and PF-C. To illustrate the performance variance between "positions of play" for FG% and REB, (see Figures 1 & 2).

Figure 1: Field goal (%) by positions

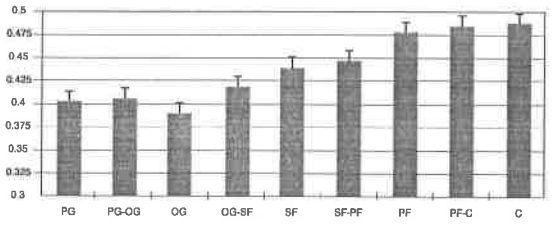
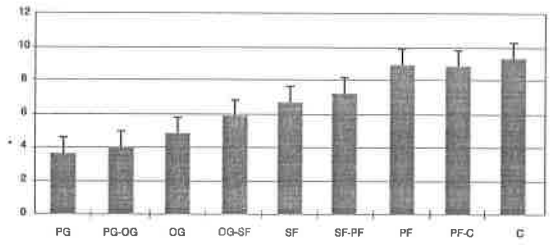


Figure 2: Rebounds by positions



For the norms established in the study (see Appendix A), six of the nine performance factors (REB, PF, TO, AST, BLK, & STL), were developed in relationship to "minutes played". For the model, a "real game" or 40 minutes of play is the common denominator for these factors.

The performance factors; FG%, 3FG%, and FT% are not adjusted for time, as the relationship of shots made to shots attempted is the common denominator for these factors.

A number of measurement features incorporated into the model add to the accuracy, fairness and flexibility of the evaluation process. The primary measurement concept that adds accuracy to the evaluation process is the "minutes played" measurement concept. Here, players are graded in relationship to the number of minutes they play. Therefore, a player who plays 20 minutes in a game and makes six rebounds would achieve a better grade on that factor than a player who plays the same position and makes six rebounds in 30 minutes of play. With this measurement concept built into the model, an accurate means to quantify differences in the quality of performance is established. Another feature incorporated into the model that adds a degree of fairness to the evaluation process is a series of "time thresholds" that must be met or the factor is not evaluated. This aspect of the measurement construct gives players who play minimal amounts of time an opportunity to be evaluated on a fair basis. The "time thresholds" are based on the average number of minutes per "real game" required to score in each performance category. The software also provides coaches the flexibility to eliminate from a player's evaluation 3FG% and BLK if the coach determines that either of these factors are not relevant to the players performance. (e.g. 3FG% for centers or BLK for point guards.)

To illustrate the results produced by the model, Figures 3 & 4 represent a single game and a season-to-date evaluation of a Division I player from the University of Minnesota in 1997.

The statistical data in the left column represents the raw game or seasonal statistics.

Figure 3: Single game player evaluation

INDIVIDUAL game STATISTICS DISPLAY			
Name: Bobby Jackson		Position: OG	
Game date: 03/22/97		Opponent: Clemson	Score: 90:84
Minutes played	40		
Field goals made	11	FG - 55.0%	
Field goals attempted	20	3.722	
Three pt.shots made	2	3FG - 50.0 %	GPS-3.689
Three pt.shots attempted	4	3.640	
Free throws made	12	FT - 92.3 %	
Free throws attempted	13	3.706	
Rebounds	9	RB - 3.874	
Personal fouls	4	PF - 1.904	
Assists	0	AS - 0.000	
Turnovers	1	TO - 3.665	
Blocked shots	0	BS - Not Graded	
Interceptions/steals	2	IS - 2.849	

Figure 4: Season-to-date player evaluation

SEASON TO DATE STATISTICS DISPLAY			
Name: Bobby Jackson Position: OG			
Game date: 07/01/97 Opponent: Seasonal (35 Games)			
Minutes played	31.5		
Field goals made	191	FG – 44.1 %	
Field goals attempted	433	2.738	
Three pt.shots made	31	3FG – 32.0 %	GPS–3.689
Three pt.shots attempted	97	2.000	
Free throws made	121	FT – 78.6 %	
Free throws attempted	154	2.930	
Rebounds	6.1	RB – 3.613	
Personal fouls	3.03	PF – 2.048	
Assists	3.97	AS – 3.665	
Turnovers	2.11	TO – 2.364	
Blocked shots	0.229	BS – 2.064	
Interceptions/steals	2	IS – 3.659	

Grades for each performance factor are displayed in the right column. The grade in the enclosed box represents the overall grade or GPS. It is important to note that the Graded Performance Score (GPS) is not merely the cumulative average of the nine Scaled Performance Scores (SPS). The GPS is calculated by scaling the cumulative-SPS to a national norm for the "position of play". In the model under study, this grade is also adjusted to reflect weighting factors incorporated into the BES-Weighted model.

The original BES model which measured player performance in relationship to "position of play" and "minutes played" was based on a common set of performance criteria. The model included a context to measure the criteria, and provided a measurement system inherent to the structure of the sport. This model however, did not take into account the importance of each individual performance factor and it's relationship to the position. Most coaches would agree, for example, that assists are more important to the role of a point guard than they are to the role of a center and that rebounds have a greater importance to a center's contribution to the team than they do to a point guard.

The purpose of the study was to develop a BES performance evaluation model that incorporates weighting factors based on the importance of the factor to the "position of play".

It was then to test the factor weighted model to determine if it significantly increased the validity of scores produced for overall performance.

Method

To collect data for a factor weighted model, a group of expert Division I head coaches throughout the nation were surveyed. One coach from each of the 32 Division I conferences was surveyed. Eighteen coaches completed surveys. Coaches were asked to judge the importance of the nine game related performance factors in relationship to position. A Likert scale ranging from 0, least important to 5, most important was used. The scores produced (see Appendix B) were then incorporated into the original BES model to produce factor weighted scores for overall performance. To demonstrate how the weighting factors were incorporated into the model and illustrate their effect on a sample player's overall grade or GPS see Figure 5.

To determine if the BES factor weighted model would increase the validity of scores produced for overall performance, a second survey was conducted. In this survey, a

Figure 5: Factor Weighted Model

Name: Rasheed Wallace College: Univerity of North Carolina Position: Center

PF	SPS		PFW		WSPS
FG%	3.733	X	4.588	=	17.127
3FG%	NG	X	0.882	=	NG
FT%	2.410	X	4.294	=	10.349
REB	2.780	X	4.889	=	13.591
PF	3.375	X	3.667	=	12.376
AST	1.872	X	2.588	=	4.845
TO	3.356	X	3.706	=	12.437
BLK	3.814	X	4.056	=	15.470
STL	1.277	X	2.167	=	2.767
	22.617		29.955		88.962
	(2.827)				(2.970)

Unweighted GPS=(2.827–2.0)/1.0=0.827 Zi=0.7967x4=**3.187**

Weighted GPS=(2.970–2.0)/1.0=0.970 Zi=0.8340x4=**3.336**

Key: Perfomance Factor (PF), Scaled Perfomance Score (SPS), Perfomance Factor Weight (PFW) and Weighted Scaled Performance Score (WSPS). No Grade (NG) indicates the factor was not evaluated.

group, (n = 15) of expert Division I coaches were randomly selected throughout the nation. Ten coaches returned completed surveys. Coaches were asked to evaluate the overall performance of a sample group of players (n = 45), five players from each "position of play" category. The identity of each player was kept in confidence. The following data was presented to coaches for evaluation: position of play, the number of games played, the average number of minutes played per game, and the seasonal

statistics for the nine performance factors graded by the model.

The seasonal statistics used for the sample of players evaluated by the coaches were randomly selected from the positional categories of players making up the national performance norms. This insured representation of a variety of performance levels in the sample group. Coaches were asked to grade overall performance on a letter grade scale of; A, A-, B+, B, B-, C+, C, C-, D+, D, D-, & F. The following grade points were assigned to each letter to establish a numerical value for criterion scores; A = 4.0, A- = 3.67, B+ = 3.33, B = 3.0, B- 2.67, C+ = 2.33, C = 2.0, C- = 1.67, D+ = 1.33, D = 1.0, D- = 0.67, & F = 0.0.

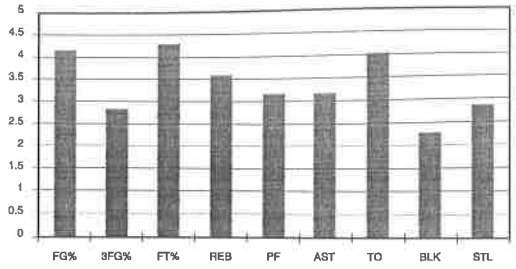
Due to the relatively large group of subjects asked to be evaluated (n = 45), coaches were paid a stipend of 20 dollars. The coaches' criterion scores were then correlated with scores produced on the same group of subjects for both weighed and unweighted BES models.

Results

To develop a generalized picture of each performance factor's overall importance to the game, positional scores from the factor weighting survey were combined. This manipulation of the data helped to separate performance factors into four categories relative to their importance to the game and their variance between positions. The categories include: (a) important-little variance (FG%, FT%, TO), (b) important-great variance (REB), (c) less important-little variance (PF, STL), and (d) less important-great variance (3FG%, AST, BLK). Categorizing the factors may help to shed light on the results of validity study below. To illustrate each performance factor's overall importance to the game see Figure 6.

As stated above, individual performance factors seem to fall into four categories. For an example of the differences between a selected factor's importance to the game, and that factor's variance between positions see Figures 7 & 8. Here, the bar graphs show that FT% is very important to all positions, whereas rebounding showed great variance in its importance to position.

Figure 6: Performance factor weights



The results of the validity study indicate that both models correlated highly with criterion scores produced by the coaches. For the BES-Unweighted model, r = .757, and for the BES-Weighted model, r = .798. To determine if there was a significant

Figure 7: Factor weights - FT%

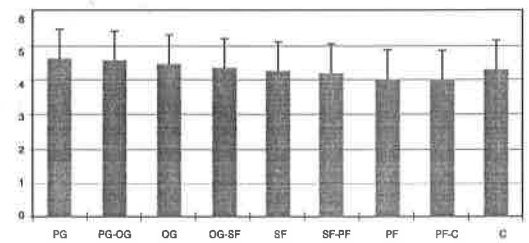
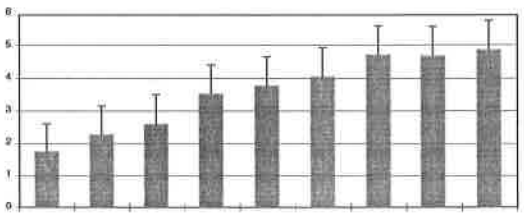


Figure 8: Factor weights - REB



difference between the correlation scores generated, a Difference in Correlation for Independent Groups one-tail test, (n = 45) was used. In order for the difference to be statistically significant, the difference between the coefficients would need to reach 0.17. The actual difference was 0.041. Therefore, factor weighting did not increase the validity of overall performance scores produced by the BES- Weighted model.

Discussion and Conclusions

The results of the factor weighting study indicate that a number of performance factors are more important to the game

than others, and that some performance factors varied greatly by position, while others did not. It is not surprising that coaches viewed FG%, FT%, REB, and TO as most important to the game. Both FG% and FT% have a great influence on the game because they directly effect point production. Rebounds and turnovers also effect point production, but in an indirect manner. Both are a means of increasing total possessions which lead to scoring opportunities. Steals also increase the number of possessions, however, so few steals are recorded in most games that they can not be viewed in the same light as rebounds and turnovers. The National Collegiate Athletic Association (NCAA) basketball statistical trend for 1997 indicate that the average number of steals per game per team averaged 7.5, rebounds 36.2, and turnovers 15.7.

Five of the nine performance factors (FG%, FT%, PF, TO, & STL), show little variance between position, whether they are viewed as important to overall play or not. The four factors that do show a high degree of variance between positions (3FG%, REB, AST, & BLK), with the exception of rebounds, are viewed as the least important factors in the judgement of coaches for their overall impact on the game. The importance of 3FG% however, may be greater to the game than the data indicates. As 3FG% is a subset of FG%, coaches may view this factor as very important, but only to select a group of players. This then would tend to hide or skew its overall importance to the game.

Appendix A: National Perfomance Norms - "real game" means and standard deviations - Positions

	PG	PG-OG	OG	OG-SF	SF	SF-PF	PF	PF-C	C
FG%	.401(.109)	.405(.075)	.390(.109)	.418(.115)	.440(.104)	.447(.097)	.478(.105)	.485(.096)	.487(.111)
3FG%	.315(.159)	.327(.136)	.320(.134)	.286(.134)	.294(.169)	.295(.196)	.240(.249)	.306(.299)	.198(.265)
FT%	.686(.159)	.679(.152)	.683(.166)	.649(.157)	.651(.154)	.644(.153)	.597(.163)	.591(.163)	.589(.164)
REB	3.559(1.267)	3.927(1.451)	4.809(2.258)	5.829(1.697)	6.654(2.238)	7.728(1.939)	8.904(2.639)	8.770(2.203)	9.361(2.886)
PF *	3.512(1.473)	3.709(1.382)	3.894(1.817)	4.103(1.522)	4.339(1.772)	4.630(1.719)	5.071(1.819)	5.274(1.857)	5.591(1.885)
AST	5.325(2.172)	4.208(1.462)	2.990(1.484)	2.831(1.233)	2.388(1.310)	1.997(1.075)	1.590(1.005)	1.507(0.847)	1.439(0.933)
TO	4.009(2.003)	3.619(1.360)	3.317(1.675)	3.139(1.138)	3.255(1.390)	3.090(1.254)	3.045(1.246)	3.009(1.081)	3.111(1.105)
BLK	0.242(0.710)	0.210(0.422)	0.269(0.519)	0.398(0.414)	0.599(0.748)	0.731(0.765)	0.872(0.904)	1.106(0.881)	1.542(1.234)
STL	1.895(1.036)	1.790(0.755)	1.526(0.844)	1.598(0.796)	1.591(1.261)	1.513(0.874)	1.219(0.684)	1.155(0.641)	0.981(0.685)

Note: For REB, PF, AST, TO, BLK, & STL, means and standard deviations were calculated for a "real game" or 40 minutes of play.

Appendix B: Factor Weighting Scores - means and standard deviations - Positions

	PG	PG-OG	OG	OG-SF	SF	SF-PF	PF	PF-C	C
FG%	3.556(.856)	4.059(.659)	4.389(.502)	4.062(.680)	3.833(.707)	4.125(.719)	4.278(.669)	4.235(1.019)	4.588(.618)
3FG%	3.727(.826)	3.882(.787)	4.444(.705)	3.812(.750)	3.333(.767)	2.375(1.025)	1.722(1.018)	1.187(1.223)	.882(1.111)
FT%	4.556(.784)	4.529(.624)	4.444(.616)	4.313(.602)	4.235(664)	4.176(.728)	4.000(.840)	4.000(1.173)	4.294(.849)
REB	1.706(.985)	2.250(.937)	2.588(.795)	3.500(.632)	3.778(.732)	4.059(.827)	4.722(.461)	4.706(.985)	4.889(.323)
PF	3.235(.930)	3.125(.957)	2.611(1.037)	3.000(.730)	2.889(.758)	3.118(.857)	3.389(.916)	3.471(.943)	3.667(1.085)
AST	4.500(.985)	4.059(.899)	3.444(.984)	3.176(.809)	3.059(.748)	2.687(.479)	2.389(.778)	2.500(.894)	2.588(1.004)
TO *	4.833(.514)	4.294(1.047)	4.222(.647)	4.059(.966)	3.824(1.015)	3.647(1.115)	3.588(1.228)	3.812(1.167)	3.706(1.213)
BLK	0.588(.870)	0.937(.929)	1.000(.907)	1.938(1.187)	2.111(1.278)	2.706(1.213)	3.444(1.338)	3.687(1.302)	4.056(1.162)
STL	3.056(1.056)	3.250(.856)	3.222(1.003)	3.294(.985)3	3.059(.899)	2.812(.911)	2.611(1.037)	2.312(1.014)	2.167(1.043)

Note: The weighting factor scale ranges from 0-5.

Since the most important performance factors to the game with a few exceptions have the least variance between positions, the difference between models was not powerful enough to cause a greater disagreement between the correlations. This may also be due to the relatively small sample size coaches were asked to evaluate.

Despite the lack of a statistically significant difference between the BES-Weighted and BES-Unweighted models, factor weighting should add to the face validity of the model. The scores generated by the model should give coaches and players a greater knowledge of results which they would not have access to otherwise. The knowledge of results produced by the Basketball Evaluation System will help players and coaches to more clearly define a player's strengths and weaknesses, as well as to act as a means to monitor a player's progress from game-to-game and from season-to-season. Coaches have often stated that it is the "little things" or the details that make the difference between winning and losing. If this is the case, then a more complete understanding of a player's performance will help to make a difference for the player and for the team.

There are many factors which influence the overall performance of players. Many of these factors cannot be easily measured. Hustle, court sense, and the ability to perform in pressure situations are some of the intangible factors that contribute to successful play. The Basketball Evaluation System makes no claim to measure the intangible aspects of the game. These aspects may best be left to the coach's knowledge, judgement and intuition. The Basketball Evaluation System does however give coaches the ability to measure the performance aspects of the game that are often overlooked because coaches do not have an objective tool by which to measure their value to a player's performance.

References

1. Burson, J.F. (1973). Developing a systematic procedure for basketball film analysis and utilization. (Unpublished doctoral dissertation), The Ohio State University.

2. Deming, W.E. (1975). The logic of evaluation. In: *Handbook of Evaluation Research,* (Struening, E.L. and Guttentag, M., Eds.), pp 1, 56. Beverly Hills, CA: Sage Publications.

3. Elbell, E.R. and F. Allen (1941). Evaluating team and individual performance in basketball. *Research Quarterly,* 12(3):538-557.

4. Swalgin, K.L. (1987). A computer assisted quantitative analysis and evaluation system of individual performance by position of play for men's division I college basketball. (Unpublished doctoral dissertation), pp 31-32, The Ohio State University. pp 31-32.

5. Swalgin, K.L. (1992). Athletic performance evaluation: a computerized model for team sport players. *Science Periodical On Research and Technology in Sport (SPORTS),* Coaching Association of Canada, Strategy GY-1; 12(5):1-8.

6. Swalgin, K.L. (1993). The relationship of the basketball evaluation system (BES) to criterion measures of performance in men's division I college basketball. *The Applied Research in Coaching and Athletics Annual,* 8:226-245.