A Ranking Model for Two Women's Team Sports

by

Drucilla Ann Rasberry

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science in Education

in

Health and Physical Education

APPROVED:

M.L. Driscoll, Chairman

R.K. Stratton

R.G. Oderwald

June 1986

Blacksburg, Virginia

A Ranking Model for Two Women's Team Sports

by

Drucilla Ann Rasberry

M.L. Driscoll, Chairman

Health and Physical Education

(ABSTRACT)

There are six women's team sports recognized by the National Collegiate Athletic Association (NCAA): field hockey, basketball, soccer, volleyball, lacrosse, and softball. Associated with each of these sports are polling procedures conducted to indicate national rankings in terms of quality of performance. These polls generally are referred to as "Top Twenty Polls."

Questionnaires were designed and mailed to the coaches and selection committee members associated with the two study sports: basketball and field hockey. Respondents were to rate a series of questions pertaining to either basketball or field hockey as to the relative importance in the ranking process of the content of each question.

Returned questionnaires were compiled and a frequency analysis done on the questions. Analysis showed there to be no difference between sports, or among divisions within a sport as to the selection methodology used to develop a Top Twenty Poll. A model was developed to provide unbiased information to the selection committees designed to allow for comparisons which might otherwise not be available. Following analysis, the following variables were chosen for the model: win/loss record against common opponents, scoring margin, strength of schedule during the season, "credible" losses, record for the week, record for the season, record against teams 1-10 in the top twenty poll, record against teams 11-20 in the top twenty poll, and a performance quality indicator statistic.

Acknowledgements

As I worked on this project I began to think of the concept of my model as a "friend," and not always as a particularly welcome one, at that. This nameless friend appeared to follow me around; at what seemed completely inappropriate times, it would creep into my consciousness, demanding my attention. I slowly began to realize applications from past classes, my work, my general environment...the learning processes I had been practicing on and off for the past 30 years were becoming techniques to be applied.

Sometime shortly after 28 January 1986, my friend gained a name. I began calling my model, indeed my entire research effort, MAC without completely realizing it. So, unofficially and affectionately, MAC it is...as a thank you to the many teachers I have been fortunate enough to have had in my life who were like Christa McAuliffe, encouraging us to seek to learn for knowledge's sake, to "reach for the stars."

There are so many people who directly or indirectly had a great deal to do with my completion of this program. They know who they are, and also that they have my deepest gratitude.

I would like to take this space to recognize the support of the following people:

Dr. M.L. Driscoll - committee chairman, gifted motivator...each time I began to slow down and hang my head, she was able to renew my enthusiasm. It is not an overstatement to say this effort would not have been completed without her support, patience, and not always

gentle prodding. Knowing which motivational methods are effective on which students is often difficult, yet Dr. D. is a master at this task.

Dr. R.K. Stratton, and Dr. R.G. Oderwald - both of whom have known me for quite some time, but agreed to serve on my committee anyway. Both shared with me a great deal of expertise, support, and a tremendous love of sports.

Ms. J.K. Kafer - primary women's administrator and long time friend. Her contacts were invaluable; without her cooperation and input, this project would never have been begun.

Dr. R.H. Giles, Jr. - teacher, employer, friend...whose influence on me will be realized for many years to come. A great deal of my personal growth of the last few years is due in part to him.

Dr. Fred Servello - my SAS expert and friend...without his knowledge of SAS and sports, and his generosity with his time, I would have missed many more of my deadlines than I did. He also helped with the control data.

Those who helped rank the control data set: Dr. Paul Angermeier, Ron Carr, Kevin Downey, Nathan Garner, Paul Lawrence, Justin Scroggs, and Dennis Vaught...sports enthusiasts, all.

Jane Scroggs - good friend and fellow miserable grad student...she helped to keep me moving forward, helped rank the control data, and served as a sounding board for frantic ideas as the spectre of my final deadlines chased me.

My family - parents Earl and Polly Rasberry, and brother Lee...their contribution cannot be fairly described in a few short words. It is difficult to express very private thoughts in a very public document, however, I owe much to these wonderful people. Thank you, and much love.

Finally, I would like to dedicate this thesis to the memory of my aunt, Margaret C. Gulledge. This very special woman encouraged me to be creative and to write. Within this effort, I have tried to do both.

In creating, the only hard thing's to begin.

A grass blade's no easier to make than an oak.

James Russell Lowell

Table of Contents

List of Tables

List of Illustrations

Chapter I

Introduction

There are six women's team sports recognized by the National Collegiate Athletic Association (NCAA): basketball, field hockey, lacrosse, soccer, softball, and volleyball. Associated with each of these sports are polling procedures which are conducted to indicate national rankings. Generally, twenty teams are ranked, with "1" being the highest rank, and "20" the lowest. These "Top Twenty Polls" are in terms of quality of performance on a weekly and/or seasonal basis.

The polls for each of the women's sports are determined by selection committees which are sanctioned by the NCAA. There are regional committees and national committees, the membership of which consist of coaches, athletic directors, and/or others knowledgeable of the sport in question. Each regional committee meets via telephone conference call weekly, usually on each Monday, during the time their sport is in competition to discuss the teams in their region, and the games played during the preceeding week. Subsequently, three to five teams are then ranked for the region.

One member from each regional committee, plus an additional member serving as chairman, form the national selection committee for each sport. This committee meets, also via a telephone conference call, usually the day following the regional committees' calls. The

teams selected by each regional committee are discussed in depth, region by region, then a consensus top twenty poll for that week is developed. This process is repeated for each week throughout the season, the purpose being to develop weekly rankings, and a means for determining the teams to attend the national post-season tournament for that sport.

Statement of the Problem

Each individual on the NCAA selection committee has a personal method for placing teams in a particular order in the polls. Which factors are considered as well as how a factor is considered relative to any other factor varies from individual to individual. The rules that provide for the formation of the committees do not extend to the decision-making process. There is no provision for standardizing what information is presented during the conference calls. In women's sports, where there is generally less money available to the programs, there is much less opportunity for teams to share common opponents across regions. This means a lack of data which can be evaluated in order to acquire a "fair" rating of all teams. In addition to this, bias can be introduced at a number of places in the decision-making processes. Individual bias can occur through means such as personality conflicts. Regional bias can be caused by "regional pride," but can also occur simply because a means for inter-regional comparison may not exist for teams without common opponents.

The methods of ranking teams used in professional sports can not be used in collegiate sports. The professional leagues deal exclusively with win/loss percentage, as in basketball, or points multiplied by wins or ties, as in ice hockey. In addition, each team plays the same number of total games, and play the same opponents as all other teams. Comparisons can be then be made by the winning percentages within the divisions for basketball, and by total points within the divisions for ice hockey. In collegiate athletics, the difficulty is in trying to compare relative quality of performance based on a varying number of games, games varying in difficulty, and a lack of meaningful common opponents, for example.

The methodology for reducing regional and/or individual bias currently is not available to selection committees. A ranking model could be a means for providing baseline information accounting for quantitative variables in an unbiased manner. Data could be analyzed regionally, or nationally, allowing for comparisons of information which might otherwise not be available. The model would dictate that each team be considered relative to a predetermined set of variables. From this standardized data, further subjective discussions could then be carried out; the result being a more unbiased comparison of teams.

To the knowledge of the investigator, there has been no previous research into this area as pertains to women's team sports. This research was undertaken to determine an inexpensive, uncomplicated method for standardization of ranking women's athletic teams, regardless of regional competition and/or team sport involved.

Purpose and Objective of the Study

The purpose of this study was the development of a model for ranking athletic teams participating in NCAA women's sports. The sports studied were basketball and field hockey. The model is adaptable to all NCAA women's team sports, with some modification of the performance variable as determined by the offensive statistics relative to a particular team sport.

The objective of this research was to determine if a method could be developed which would increase the use of computer technology in establishing continuity in ranking techniques used within women's team sports. The specific objectives were: (a) to determine the methodologies used by the individuals serving on the selection committees, (b) to determine whether selection methodologies differed from sport to sport, (c) to determine whether these methodologies differed by Divisions I, II, or III within a sport, and (d) to develop a model to aid in creating top twenty polls.

Summary

In the preceeding pages, a need was established for standardizing the techniques used for ranking NCAA women's team sports. Through established selection committees, each women's team sport conducts polls for the purpose of ranking teams in terms of quality of performance relative to each other team. Currently, there is not a method to allow for comparisons based on purely statistical, or objective, data which may compensate for, or significantly reduce, regional and/or individual bias.

Chapter II

Review of the Literature

Introduction

Specific studies as to ranking models for women's team sports have not been made. However, studies have been made in terms of statistical forecasting, prediction, decision theory, and simulation and modelling. There is some related work being done in modelling pertaining to athletics.

The scope of this research was not to develop a model to be used for prediction as such. That is, it will not be used to take a known ranking for week X and from that, predict the ranking for week Y. Instead, the function of the algorithm is to assimilate known data for some pre-selected set of teams and rank those teams based on common selection criteria. While this model is not a forecasting model in the "traditional" sense, many of the same strategies are incorporated herein, making the forecasting and prediction literature relevant.

Forecasting and Prediction

Forecasting and *prediction* are nearly synonymous terms. To predict, according to *Webster's New Collegiate Dictionary* (1981), is "to declare in advance, especially to foretell on the basis of observation, experience, or scientific reason." To forecast is "to calculate or predict...usually as a result of rational study and analysis of available pertinent data; to indicate as likely to occur." Traditional prediction or forecasting theory is based on modelling to simulate specific conditions, ranging from, for example, aspects such as commercial pricing strategies to war games.

Forecasting can be generally divided into two broad categories: quantitative and qualitative. Qualitative methods are largely intuitive and are not necessarily dependent on previous data. Usually, these methods are not reproducible due to a lack of an explicit method for incorporating the information into the forecast (Abraham & Ledolter, 1983).

Quantitative models are stochastic, or statistical, and are reproducible (Abraham & Ledolter, 1983). The predictive model is measureable, thus quantitative, and is based on data collected for past events. The basis for this statistical prediction is that the analysis involves two "experiments," *x* and *y*. If x is a variable about which information is gathered in order to predict some future condition of y, then there must be a link between x and y (Aitchison & Dunsmore, 1975).

In a quantitative forecasting system, there are two main components: the model building stage and the forecasting stage. Figure 1 on page 7, adapted from Abraham and Ledolter (1983), shows a conceptual flow chart of this type of system. No model is absolute, so an important part of the model building process is checking. This step "involves looking for model inadequacies or for areas where simplification can take place" (Jenkins, 1976). The checking step serves as a means of evaluating the reasonableness of the model, fine tuning and refining to improve the accuracy of the forecasts.

Table 1 on page 8 contains guidelines presented by Jenkins (1976) which can be used in building forecasting models. The fundamental premise is that the model builder under-

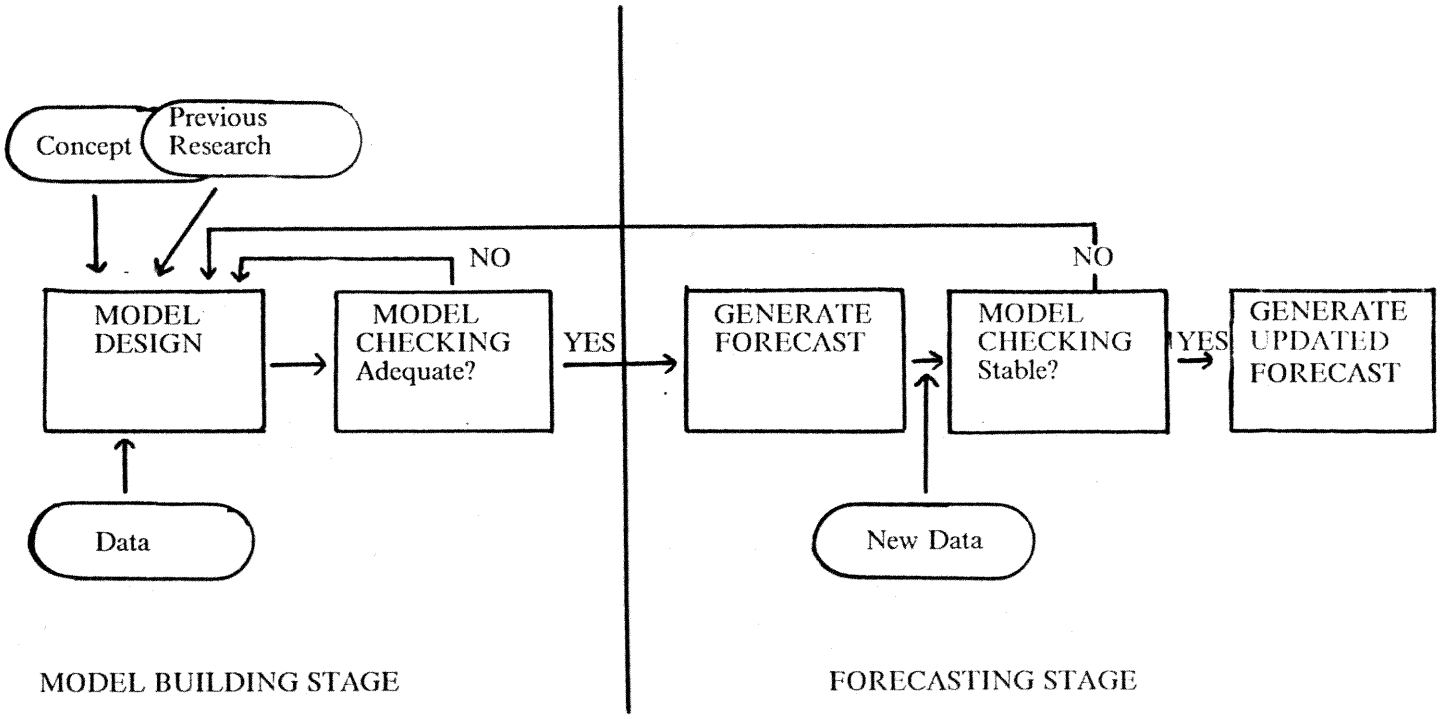

Figure 1. **Conceptual flow chart of a quantitative forecasting system:** adapted from Abraham and Ledolter (1983).

Table 1. Jenkins' guidelines for building forecasting models.

1. Understand the problem and the purpose of building the model.
2. Understand the decision-taking system which the model will serve.
3. Work out early on how the model is to be implemented.
4. Structure the quantitative model by building a conceptual model of the appropriate environmental system, displaying the mechanisms involved.
5. Select the data carefully, understand its limitations and plot it in a variety of ways.
6. Aim for simple models, involving few variables, first and then later, if necessary.
7. Proceed iteratively via
   -Identification (Specification)
   -Estimation (Fitting)
   -Checking (Criticism)
8. Aim for parsimony in parameterisation - avoid over parameterisation.
9. Understand what the model has to say about the data.
10. Conduct experiments with the model (simulations) to inderstand its limitations.
11. Present the results from the model in simple terms to those who have to use it.

stands the problem to be solved. Without this understanding the model may be irrelevant (Jenkins). Understanding the decision-making system of which the model will be one part is a necessary step. This is related to understanding the problem and the objectives of the model. The scope of the overall decision-making system influences the role of the model and how the informational output from that model is used.

The principle of parsimony is an important concept to which to adhere when building models. Also known as Ockham's razor, this principle states that in a choice among competing models, other things being equal, the simplest is preferable (Abraham & Ledolter, 1983). Not only are simpler models usually easier to manage and interpret, but estimation of every additional parameter increases the variance of the error by a factor of 1/n (see Abraham & Ledolter).

Suitable data is also necessary. Building a variable into a model for which data cannot be readily obtained may make the model too costly or difficult to implement, rendering it virtually worthless.

Predicting the performance of a future condition of y depends in part on the consequences of being wrong, or the risk factor. It may be enough to be "close" to the outcome, or it may be necessary to determine the consequence of being "badly" wrong. If these consequences can be quantified, then the problem can be presented "as one of statistical decision theory" (Aitchison & Dunsmore, 1975).

Decision Theory

In the context of this study, decision theory is a term describing the processes involved in making decisions. It is relevant not only to design and structure in model building and simulation, but also to general research behavior and systems evaluation. Models are tools which are successful only if the application is relevant to a successful system.

Figure 2 on page 11 is from Giles (1986), and demonstrates his *General System*. A general system model is one that is used to aid in objectives analysis. Sirgy and Giles (1985) outline seven components of this system: 1) objectives, 2) inputs, 3) processes, 4) outputs, 5) comparator, 6) feedback, and 7) feedforward; adding "outputs" and "comparator" to Giles' figure. Extensive definitions may be found in Sirgy and Giles, briefly: *objectives* are "a priori" decisions about expected results of a system, or the dependent variable(s); *outputs* are analyzed results of the system; *inputs* are the independent variables; *processes* are thought of as the interactions among or between the variables; *comparator* compares outputs to objectives; *feedback* is an operation, "controlling and regulating" the system in order to achieve the objectives; and *feedforward* is a design element, a special type of input used to shape the model toward a future desired outcome (Sirgy & Giles).

Sirgy and Giles (1985) also identified seven orders of objectives. The hierarchy follows the numerical listing of the general systems components found above. As discussed in the previous section, the objectives identify the desired conditions, the goals and purposes for the system or model. The feedback objectives help in adjusting the model to yield outcomes more closely related to the desired or expected values.

One means of formulating a decision as a general system so that the power of system theory can be utilized is to list objectives and alternative actions or potential decisions. Then the objectives are weighted or values assigned which are "expressions of relative worth or significance" (Giles, 1986) to the objectives. This can be done by utilizing an objectives-weighting matrix. The procedure then involves determining the product of each weighted objective and the effectiveness of each alternative in achieving that particular objective. Normalized effectiveness scales are compiled in a matrix. Multiplying the weight by the effectiveness results in a weighted effectiveness for each possible action or decision alternative. The "best" course of action or alternative is then identified as the one with the highest sum of these products (See Giles).

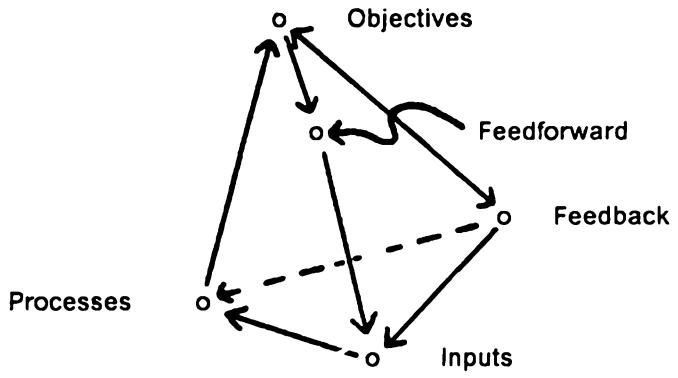

Figure 2. The General System: from Giles (1986).

Questionnaires

Questionnaire design is discussed in Dillman (1978). Various examples of different designs, target populations, and mailing techniques are discussed thoroughly in the book. Response rate is defined by Dillman as the number of returned questionnaires divided by the sample size, or the number of questionnaires mailed, expressed as a percentage. Response rate can be used to determine whether bias exists in the returned questionnaire sample be it in terms of geographical bias, such as returns skewed toward one region or another, or divisional bias, such as significantly greater returns from one division in a sport over those from the other divisions.

Statistical Methods

Specific statistical techniques relative to studies in the athletics field are discussed in the following section. Some general statistical methods and theories are discussed in this section.

Regression analysis is generally used to examine the relationship among variables (Whittle, 1983; Mosteller & Tukey, 1977; Kerlinger, 1973). The regression is used to explain the relationship between a response, or dependent, variable and a set of explanatory, or independent, variables (Abraham & Ledolter, 1983). To illustrate this, Abraham and Ledolter used the following example: forecasting monthly beer sales, the dependent variable, from the price of beer, the target consumer group's disposable income, and seasonal temperature, the independent variables.

A condition of which to be aware when using regression analysis is that of multicollinearity. Multicollinearity occurs when several of the independent variables are highly correlated, thus making the matrix nearly linearly dependent, and is common for models fitted on observational data (Abraham & Ledolter, 1983). When one variable can be ex-

pressed exactly, or approximately, by a linear combination of the variables, then a degree of collinearity exists. If multicollinearity is severe, there is difficulty in determining the effect attributed to each independent variable (Abraham & Ledolter; Ott, 1977), as well as in determining how much of the variance is explained by each variable.

One means for simplifying comparisons is to use indices. An index is an observable phenomenon which is substituted for a less observable phenomenon; more generally, a number that is a composite of two or more other numbers (Kerlinger, 1973). Examples of indices include averages, sums, and correlation coefficients.

In addition to simplifying comparisons, use of indices may also allow comparisons that could not otherwise be made, or possibly could be made only with extreme difficulty (Kerlinger, 1973). An index that is meaningful is one expressed as a quotient: x/y, ratios or proportions. A percentage is a variation of a proportion.

One of the more useful forms of the quotient index is that expressed with a range of between 0 and $\pm$ 1.00. Using this range tends to make the index independent of numbers of observations and aids in comparison from sample to sample, and study to study (Kerlinger, 1973).

Athletics

While no ranking model has been developed for women's team sports, much relevant research has been conducted in athletics. The following is a discussion of some of this research, much of which contributed ideas and insights to different aspects of this model.

Bennett and Flueck (1983) summarized offensive performance models used in baseball. The most obvious offensive performance evaluation is in terms of runs scored, however, the statistic most often used for looking at offensive performance is team and individual batting average. Ten performance models were discussed as well as the run production models associated with each. The ten estimators are: batting average, slugging percentage, on base

average, offensive average, run production average, offensive performance average, the Cook Scoring Index, the Modified Cook Scoring Index, the D'Esopo-Lefkowitz Scoring Index, and the expected run production average. Some of these are complicated equations, but they are all indices of individual offensive performance. Combining individual performance values gives an index for team performance which in turn can be used for comparison among teams.

An example of using weighted objectives to find optimal solutions is found in a program that uses computer-based assignment and matching techniques to schedule referees to collegiate football games and umpires for major league baseball games (Evans, Hebert, & Deckro, 1984). The model evaluated elements such as distance to future contests, number of assignments already made to an official or crew, and airline schedules, in scheduling the officials. The algorithm ranks the 10 best assignments in order of increasing cost, allowing for "human interaction" in the decision process.

Using probablility to predict the outcome of a sporting event is demonstrated by Sawusch and Summers (1984). The equation used to express the probability that a stronger team will win a seven game series is:

$$p^4 + 4p^4q + 10p^3q^2x + 10p^2q^2(1-x)p$$

where p is the probability that the stronger team will win ($\geq .5$), q is the probability that the weaker team will win (1-p), and x is a conditional probability particular to a given sport. This information is passed as parameters to a computer program which then "plays" 1000 games to predict the outcome of the series.

Haberman (1974) used a study to show how quantal-response models may be used to determine probable standings for American League Baseball teams in 1948. The model used paired comparisons to evaluate probabilities that "stronger" teams would win against "weaker" teams.

A means for determining the present value of a player has been developed. This methodology uses required inputs to determine a player's worth in terms of salary. For ex-

ample, the required inputs for a pitcher are: number of wins, incremental home attendance when he pitches, number of home games he pitches, incremental home attendance when he does not pitch, probability of winning the division title, probability of winning the pennant, probability that the division title is due to the pitcher, probability that the pennant is due to the pitcher, attendance at playoff and World Series games, number of those games, and the incremental salary increase of other players due to the pitcher (Mitchell & Michel, 1977).

Price and Rao (1977) developed a model for evaluating performance of professional basketball players. This model used regression analysis to examine team characteristics and determine individual player contributions to the team's relative success.

Trend analysis relative to football was conducted by Ryan, Francia, and Strawser (1977). They used frequency analysis to demonstrate tendencies for teams and their opponents. They found that weighting would be an attempt "to provide a standardization of all data accumulated on a specific team against all of its opponents," however, a grosser level of reliable data could be obtained without weighting.

The concept of adjusting standings for strength of teams played was addressed by Smith (1977). Playing multiple games against weak opponents where each game is weighted equally does not provide for "fair" comparisons with teams who have played stronger opponents. One means for making fairer comparisons is to adjust the standings in terms of opponents' strength. The probability that team *i* will win each game played against team *j* is estimated for each pair of teams based on games already played. Since this probability is estimated based on games already played, sufficient games must have been played for this to be meaningful.

In an article in the *NCAA NEWS* (1986), Van Valkenburg discusses the computer program the NCAA uses to rank difficulty of schedules. The program averages the winning percentages of all Division I opponents to get rankings for Division I schedules; wins against non-Division I opponents are not counted. In comparison with wire service Top twenty polls, Van Valkenburg found 3 teams had a schedule ranked below 50th place, and 4 or more others had schedules ranked considerably lower.

Rosner (1976) developed a series of probability models as an analysis of professional football scores. One model examined the overall frequency distribution for the number of scores using an offensive effect, points scored, and a defensive effect, points allowed. Rosner found "both the offense of the team in question and the defense of the opposing team have a significant effect [on score] while the home/away factor does not have a significant effect."

A method for ranking teams using collegiate football as the application was developed by Leake (1976). This method has its roots in electrical networking theory. Five numerical indices were assigned each team:

1. *Team index*: an arbitrary integer assignment used for computer processing;
2. *Team rating*: based on results of outcomes of games, an index is assigned each team. The highest index is assigned a value of zero, which is the reference to which all other ratings are computed;
3. *Team rank*: rank based on the rating. For a set of size K, the team with the highest rating would receive the rank of 1 while the team with the lowest rating would receive the rank of k;
4. *Team performance*: based only on scores of games; does not take into account difficulty of schedule; and
5. *Schedule difficulty*: an average of the ratings of the teams played.

Leake identified two basic problems associated with accurate, unbiased ratings for football: subjectivity and intransitivity. Subjectivity is related to "personal biases and *a priori* assumptions of the rater or rating mechanism." Transitivity refers to the mathematical equality relationship: if A = B, and B = C, then C = A. Intransitivity occurs when A beats B, and B beats C, but C beats A. In Leake's work, intransitivity creates discrepancies between the differences in the ratings and the scores.

The point of the model is to determine objective ratings, but it requires 3 subjective criteria: distance measured between team nodes in the network graph, which in turn is related to assignment of an importance weight to each game, and the game outcome measure. One experimental measure used was 100 plus the point spread, placing the emphasis on the

win rather than the scoring margin. In effect, this lessens the severity of the point spread, and "smoothes" out games in which scores may have been run up.

Frequently, there is difficulty in recognizing a preferred element, or team ranking, or in ranking all teams when only paired comparison is known between different elements (Ushakov, 1976). This is further complicated when the evaluation differs "by one expert opinion or nonquantitative evaluation" methods. Ushakov (1976) uses networking and paired comparisons to build matrices similar to Giles' (1986) objective-weighting matrix, which are used to solve the ranking process when intransitivity is present.

Summary

Much work has been done in forecasting and prediction. The models most relevant to this study are quantitative, and are built with an understanding of the problem and objectives. Decision theory addresses the processes involved in making decisions. Evaluation and re-evaluation of the objectives are crucial steps in formulating the model. Checking of the model is done periodically through feedback, and then through feedforward. In athletics, some work has been done in evaluation and prediction. Using offensive performance estimators as an indicator of team performance can be used for team comparisons. In addition, weighting objectives has been used in assignment of officials and in ranking teams based on relative strengths.

Chapter III

Methodology

Introduction

This chapter includes a discussion on the procedure used in this study. Data were collected for two women's team sports: field hockey and basketball. These two sports were chosen because they potentially provided diversified variables for the model with one being an indoor sport and the other an outdoor sport. Examination of the diversity was important in that factors such as field conditions and temperature may have been important in the outcome of a field hockey game, but probably would not influence a basketball game. Topics discussed in this chapter are procedure and questionnaire, population, data procedure, and data analysis. Special terms used throughout are defined in the glossary found in Appendix A.

Procedure and Questionnaire

Conference calls made by the national selection committee for women's Division I field hockey provided information as to the manner in which the top 20 women's field hockey teams

were chosen. These calls were monitored anonymously, and notes were taken on the discussion pertaining to selection of the poll. The questionnaires (Appendices B, C) used in this study are a reflection of these discussions. The questionnaires were designed from points raised by members of the selection committee, as well as from the statistics collected on the NCAA official score sheets for each sport. The questionnaires were structured such that the respondents were to rate a series of questions pertaining to either basketball or field hockey as to the relative importance of the content of each question in the ranking process. The specific emphasis placed by the selection committee members is approximately the same demonstrated in the content of the questionnaires. The same questions were asked week to week, regardless of region affiliation. This demonstrated an overall effort by the committee to attempt to prevent bias from being introduced into the ranking process. An example of questions from the basketball questionnaire is shown in Table 2 on page 20.

Each questionnaire was coded to determine which questionnaires were not returned. The questionnaires were numbered consecutively within a sport; 1 through 751 for basketball, 1 through 252 for field hockey. Respondents were assured of complete confidentiality regarding participation and response.

Population

The NCAA recognizes three Divisions; each of the member institutions belongs to Division I, II, or III for each sport. In women's team sports, the exception is field hockey which has only Divisions I and III.

Mailing labels were obtained in the spring of 1985 from the Virginia Tech Athletic Association with the addresses of the coaches of each institution participating in at least one of the study sports for Divisions I, II, and/or III. The mailing labels were placed in a box, drawn in no particular order, and assigned to the questionnaires. The questionnaires were mailed to each of these coaches, along with a stamped, self-addressed envelope, and a cover letter

Table 2. Sample from the basketball questionnaire

How important is: (choose only one)	NC	NI	SI	I	VI[1]
1. Comparing records against common opponents?	1	2	3	4	5
2. Comparing scores of games against common opponents?	1	2	3	4	5
3. Scoring margins? (e.g. is 85-70 as good or bad as 50-35?)	1	2	3	4	5
4. Strength-of-schedule, beginning of season?	1	2	3	4	5
5. Strength-of-schedule as opponents on that schedule move in or out of the "top twenty?" (Does strength-of-schedule change during the season?)	1	2	3	4	5
6. Games won or lost against "top twenty" opponents?	1	2	3	4	5
7. Credible losses?	1	2	3	4	5
8. Credible wins?	1	2	3	4	5
9. Considering only the current week's win/loss record, or that of the current and the immediately previous week?	1	2	3	4	5
10. Considering the complete win/loss record up to the date of consideration?	1	2	3	4	5

[1]NC = Not Considered, NI = Not Important, SI = Somewhat Important, I = Important, and VI = Very Important.

(Appendix D) which described the study, and requested their cooperation in responding to the questionnaires. In addition, questionnaires, envelopes, and cover letters were sent to members of the regional and national selection committes for the 1984-1985 seasons. The general procedure used to formulate the cover letters and the questionnaires was based on Dillman (1978).

Data Procedure

The questionnaire results were coded and entered into a personal computer data base management system. The computer was used to manage the data and facilitate data analysis. A manual file was kept in which was recorded questionnaire number, school name, state in which school was located, whether that questionnaire was returned, the school's division membership, and any relevant comments. This served as a checklist for recordkeeping.

Separate data bases were established for each sport. In addition, duplicate data bases were set up within each sport, and each questionnaire was entered into each data base. The data were sorted into sets by division within a sport, for example, the data for Division I in basketball was in a set named DIV1 in the basketball data base. Following this, the data from each data base was compared, questionnaire by questionnaire. This was done as a means of eliminating data entry errors from the coded questionnaires. The data could then be output in a number of ways for comparison: by the total records per sport, to compare basketball versus field hockey, and by divisions to compare across divisions within a sport or by divisions across sports. Table 3 on page 22 is a matrix which illustrates this concept.

As shown in Table 2 on page 20, the valid responses for the majority of the questions were *not considered* (NC), *not important* (NI), *somewhat important* (SI), *important* (I), or *very important* (VI). Each sport had a section containing true/false questions. Respondents were instructed to select only one choice for each of the questions. Non-response by question was

Table 3. Comparison matrix.

	Field Hockey	Basketball
DIV 1	•	•
DIV 2		•
DIV 3	•	•

across sports

by division

by division within a sport

a design indicator that a particular question was not completely understood and perhaps should be eliminated from consideration.

Data Analysis

Non-response bias and potential regional bias, by division within each sport was examined. Using the NCAA handbooks for each sport, the different school-region affiliations were mapped by division within a sport. Regional response was compared to the total number of questionnaires mailed per region to determine the percent of the total which were returned. These were then compared within a division to evaluate potential regional response. These data are found in Appendices E,F; an example is shown in Table 4 on page 24.

Non-response bias deals with the question of whether the people who did not respond to the questionnaire have the same attitudes as those who did respond. In this study, attitudes were not directly addressed, but rather each questionnaire is considered to be a method by which a respondent would rank athletic teams. Additional returned questionnaires would only increase the number of individual methods of ranking. As a result, non-response bias was determined to not have an influence on this study, and no further attempts were made to retrieve the unreturned questionnaires.

A frequency analysis was implemented using the Statistical Analysis System (SAS), from the SAS Institute, Cary, North Carolina, which is available at Virginia Tech. The data were first placed into computer files by division within a sport. The data files consisted of 8 columns of information: division (DIV), question number (Q.NO), value 1 (V1), value 2 (V2), value 3 (V3), value 4 (V4), value 5 (V5), and value 9 (V9). Values 1-5 correspond to the number assigned to each valid response in the questionnaire; V9 was assigned for those questions with invalid responses such as multiple answers for a question, or no answer. Each of these "value" columns contained a number corresponding to the number of respondents selecting

Table 4. Example of questionnaire response by region.

Field Hockey - Division 1

Region	Total	Returned	% Returned
New England	17	7	41.18
Mideast	19	7	36.84
Mid-Atlantic	10	5	50.00
South	19	8	42.11
Midwest	22	11	50.00
West	5	2	40.00
Totals	92	40	43.48

that value for the specific question. An example from the basketball, Division 1 data file is shown in Table 5 on page 26.

A program was written to check the data files for data entry errors. This program read in the columns and summed them for each question. The value SUM, then, was equal to the number of returned questionnaires for each division.

A second program was written to take the data files, read in the columns, and multiply V2 by 2, V3 by 3, V4 by 4, and V5 by 5, creating columns STWO, STHREE, SFOUR, and SFIVE, respectively. A new column, RANK, was generated by summation of V1 and the "S" columns:

$$RANK = ONE + STWO + STHREE + SFOUR + SFIVE$$

The true/false questions were structured such that questions responded to with *true* were assigned a value of 4, and those with a response of *false* were assigned a value of 2. The data dealing with V9, the invalid responses, were ignored; in effect, V9 was multiplied by zero. The data file was then sorted in descending order by rank, and a new file was printed with data for division, rank, and question number.

The new data files were plotted on a horizontal bar graph. The intervals plotted were from 0 to 1200, with midpoints incremented by 5. This large range was necessary only for the total data sets; none of the individual divisions had total values this large. However, all data sets were plotted at once, requiring range values set large enough to allow for plotting the largest data set. Frequency-percentage tables, found in Appendices G,H, were generated with the bar graphs. These tables were used to locate the cumulative percentage values associated with each interval midpoint.

This analysis was used to indicate the point at which to begin to evaluate the selection criteria. Frequency bar graphs were plotted for the total samples from each sport, and for the sample split by each division within a sport. Initially, the criteria which fell closest to a 50 percent break in the data were selected. The first logical break in the data at the highest end of the curve was then evaluated. Subsequently, these were rejected for being too coarse, and

Table 5. Sample of data from basketball, Division 1 data file.

DIV	Q.NO	V1	V2	V3	V4	V5	V9
B1	Q1	0	6	22	43	35	1
B1	Q2	3	13	38	44	9	0
B1	Q3	7	25	53	19	3	0
B1	Q4	0	4	28	42	32	1
B1	Q5	1	6	18	52	30	0
B1	Q6	0	0	10	52	45	0
B1	Q7	0	0	36	53	18	0
B1	Q8	14	26	36	26	3	2
B1	Q9	0	0	28	50	29	0
B1	Q10	0	3	18	62	24	0
B1	Q11	1	1	13	45	46	1
B1	Q12	3	7	21	50	26	0
B1	Q13	1	12	41	40	12	1
B1	Q14	7	20	49	23	7	1
B1	Q15	57	31	13	3	2	1

too fine, respectively. The logical break in the data that fell closest to the upper 20 percent, or 80 percent cumlulative, was chosen as the point at which to begin evaluating the variables. The logical data breaks proved to be too uneven using this method. The decision was made to evaluate the data which fell in the upper 20, 15, 10 and percent of the distribution, disregarding any logical breaks in the curve. This analysis selected a set of questions for each sport, and for each division within each sport. These questions then became the first variables to be evaluated for use in the model. The complete ranking for each question is shown by division in Appendices I and J.

Control Data

Two data sets were created for the purpose of testing the relationships of the variables to each other, and the validity of using these variables to aid in creating a top twenty poll. Using data from professional sports (Friedman, 1978) and the collegiate ranks (NCAA NEWS), reasonable data sets were constructed. One set was to approximate statistics during the season, Table 6 on page 28, the second represented final statistics for a season Table 7 on page 29. The data were structured so that it was expected that the team ranks would not remain the same from one set of data to the other. This was done to determine whether the changes in the input data could be interpreted by the model, and also as part of the model checking and evaluation stage. The model would be used to generate a poll from one set of data, then input new data, the other data set, and generate an updated poll.

The data sets were constructed from separate sources because of an inability to locate data for the 11 variables from one source. The information published in the NCAA News generally lists the top sixteen schools in categories such as scoring margin and field goal percentage. The top schools are usually not the same in each of the statistical categories, and differ again from those ranked in the top twenty polls. Because of this difficulty in estab-

Table 6. Data set used to evaluate model - mid-season data.

Team	Rec. vs Com. Opp.	Scores vs Com. Opp.	Strength of Sch. Begin. Season	Strength of Sch. During Season	Rec. vs Top 20	Cred. Loss	Week Rec.	Seas. Rec.	Rec. vs 1-10	Rec. vs 11-20	Off. Stat.
Derring	8-0	25.9	10	9	8-0	0	3-0	20-0	4-0	6-0	+.036
Davidson	6-2	13.2	1	3	6-2	2	2-1	13-7	4-1	1-1	-.010
Williams	4-4	10.1	8	8	4-4	0	3-0	12-8	1-0	3-4	-.021
Pamplin	3-3	9.8	22	20	3-3	0	3-0	13-7	0-0	0-0	-.028
Norris	7-1	13.3	35	38	7-1	0	3-0	18-2	0-0	0-1	+.028
Whittemore	4-0	13.5	15	15	4-0	0	2-0	16-4	0-0	4-0	+.000
Patton	3-3	13.0	40	41	3-3	0	1-2	15-5	1-3	2-0	-.010
Cowgill	4-0	13.6	56	54	4-0	0	2-0	16-5	2-0	2-0	+.004
McBryde	4-3	11.9	9	6	4-3	2	3-1	13-6	1-3	4-0	-.022
Price	3-0	10.1	27	27	3-0	1	2-0	11-9	1-0	2-0	-.026
Sandy	6-2	15.0	31	31	6-2	0	3-1	16-4	3-0	3-2	+.018
Smyth	3-2	12.9	30	30	3-2	0	1-2	14-5	0-1	3-1	-.006
Henderson	3-0	13.0	41	40	3-0	2	2-0	13-7	0-0	3-0	+.010
Wallace	4-2	10.3	38	39	4-2	2	2-2	13-7	2-0	2-2	+.001
Robeson	3-4	12.2	52	50	3-4	1	2-2	14-6	3-2	2-2	+.006
Randolph	7-1	15.1	2	1	7-1	0	3-0	18-2	0-1	6-0	+.038
Holden	4-0	18.8	5	5	4-0	3	1-0	19-1	0-0	4-0	+.018
Agnew	3-0	14.8	29	29	3-0	1	2-0	16-3	0-0	3-0	+.028
Hutcheson	4-1	14.6	15	19	4-1	0	3-1	15-5	4-1	0-0	+.017
Seitz	3-2	13.0	75	75	3-2	1	2-2	15-4	3-2	0-0	-.005

Table 7. **Data set used to evaluate model - end of season data.**

Team	Rec. vs Com. Opp.	Scores vs Com. Opp.	Strength of Sch. Begin. Season	Strength of Sch. During Season	Rec. vs Top 20	Cred. Loss	Week Rec.	Seas. Rec.	Rec. vs 1-10	Rec. vs 11-20	Off. Stat.
Derring	18-8	15.1	10	9	18-8	0	3-0	33-20	8-3	10-5	+ .038
Davidson	6-17	12.9	1	3	6-17	6	1-2	26-27	5-9	1-8	-.006
Williams	7-14	11.8	8	8	7-14	2	1-2	24-29	1-8	6-6	+ .010
Pamplin	3-3	7.9	22	20	3-3	1	4-1	19-33	0-1	3-2	-.026
Norris	12-2	14.8	35	38	12-2	1	3-1	30-22	0-0	0-4	+ .007
Whittemore	14-3	14.6	15	15	14-3	0	2-1	29-23	7-0	7-3	+ .017
Patton	13-7	13.5	40	41	13-7	2	3-0	28-24	3-4	10-3	+ .000
Cowgill	11-2	13.2	56	54	11-2	0	2-1	29-25	8-2	3-0	-.010
McBryde	8-11	10.3	9	6	8-11	5	2-1	24-30	2-3	6-8	+ .001
Price	13-2	10.1	27	27	13-2	1	1-2	22-34	8-2	5-0	-.021
Sandy	6-2	18.8	31	31	6-2	0	2-2	34-19	3-0	3-2	+ .018
Smyth	6-4	15.0	30	30	6-4	1	0-3	32-23	0-1	6-3	+ .018
Henderson	6-10	13.0	41	40	6-10	4	1-2	28-28	3-4	3-6	-.010
Wallace	14-14	11.9	38	39	14-14	6	2-1	25-30	6-7	8-7	-.022
Robeson	7-14	9.8	52	50	7-14	3	1-2	23-32	5-8	2-6	-.028
Randolph	10-8	20.9	2	1	10-8	0	2-2	35-19	0-6	10-2	+ .036
Holden	4-16	15.3	5	5	4-16	7	0-3	35-21	0-8	4-8	+ .028
Agnew	3-9	13.6	29	29	3-9	2	2-1	30-24	0-0	3-9	+ .004
Hutcheson	14-7	13.0	15	19	14-7	0	3-0	29-26	8-5	6-2	-.005
Seitz	13-3	12.2	75	75	13-3	1	0-3	25-27	11-3	2-0	+ .006

lishing a real profile for twenty teams, data from the National Basketball Association (NBA) 1977 season were used.

Every attempt to reduce bias in creating the data set was taken. To avoid preconceived associations with school or team names, names of buildings on the Virginia Tech campus were used. Only academic building names were used, and those in which the evaluators had offices or classes were avoided. The team names were listed in no discernible order. The data, while from different sources, were "real" data. Data were taken from the 1977 NBA, and the NCAA statistics from two points in the 1985-1986 collegiate season. The assumption was made that data for which the rankings were known would be a more realistic test of the model than data which had been completely fabricated and thus had greater potential for ambiguities. The data values were assigned to the teams in order of team name, that is, no attempt was made to assign data values specifically to teams.

The assumption was made that the only common opponents among the teams were from within this group. The scoring margin, then was expressed in terms of common opponents. Strength of schedule was a value based on a total of 100 schedules; a value of 14 in the cell meant that Team X had the 14th toughest schedule out of 100 possible ratings. The offensive statistic was a value representing the difference in field goal percentage made and field goal percentage allowed, expressed in thousandths of a point. For example, Team X had a field goal percentage of **.481**, and had allowed a field goal percentage of **.441**. The difference, **.040**, was the value in the cell. A negative value meant Team Y gave up more percentage points than it made. The decision to use the expression of the difference between field goal percentage for and against was made after evaluating the NBA data. The NBA teams were ranked in order of winning percentage to obtain a ranking from 1 to 22. The two teams with the lowest winning percentages were dropped, as only 20 were needed. When the field goal percentage made (FG% +) was ranked, large discrepancies were detected between the two rank orders. If it is assumed that the desired variable is actually an index of performance, or overall quality, then FG% + is not an accurate measurement as such. For example, a team may have the "best" FG% + at +0.490, but may have only a .500 winning percentage as a

result of having a field goal percentage allowed (FG%-) of some inflated value. In view of this, it was decided that a better measurement of overall performance would be the difference between FG%+ and FG%-. This would give an index of relative quality based on how well the offense *and* the defense performed.

The data sets were given to 14 individuals along with instructions describing the procedure to be used in evaluating the information. Table 6 on page 28 and Table 7 on page 29 illustrate these data sets; the instructions may be found in Appendix K. The evaluators were instructed to rank the teams as if they were deciding on a Top Twenty Poll from the available information. The rankings were returned from nine of the evaluators. To obtain a "concensus" poll, the individual rankings were compiled, and numerical values assigned to each position in reverse order. For example, a team assigned to the top ranking, or position "1" received 20 points, position "2" received 19, and so on, with position "20" receiving 1 point. These were then summed to determine the consensus poll.

A correlation matrix of the data set variables was created to examine the relationships of the variables in the control data set to each other. This was done to look for potential collinearity, and to see if variable substitution would have an effect on the model. Stepwise regressions were performed to examine the relationships of the variables in models of varying sizes.

Fifteen versions of the equation were developed with substitutions of some of the 11 variables. These were then tested against the control data set to see how well each ordered the teams relative to that which the evaluators had chosen. In each equation, the variables were weighted equally, that is, each coefficient was equal to 1. From this the final model was developed.

Chapter IV

Results

Introduction

The results of the steps outlined in Chapter III are discussed in this chapter. These include questionnaires and respondents, regional bias, data procedure, data analysis, and the final model.

Questionnaires and Respondents

Two questionnaires were developed for this study as a direct result of conference calls monitored anonymously during the 1984-1985 season. The questionnaires for each sport followed a similar format. The questionnaire for basketball contained 41 questions. Responses for 2 of the questions were either *true* or *false*, while the remainder were either *not considered*, *not important*, *somewhat important*, *important*, or *very important*. The questionnaire for field hockey contained 32 questions, of which 5 were to have been answered *true* or *false*.

The questionnaires were divided into 4 sections. Section I addressed the relative importance of a particular factor considered during discussion of a poll. Section II contained the "true/false" statements. The questions in Section III were taken from the statistics kept on NCAA score sheets during each game. Section IV was used to indicate which position(s) applied to each respondent. The first 15 questions on each questionnaire, and an additional 4 questions were directly related; field hockey had additional questions relating to tied contests that did not apply to basketball, and most questions regarding the statistical section could not be directly correlated between the two sports.

One thousand and three questionnaires were mailed to coaches and selection committee members. Of those, 751 were to those affiliated with basketball, 252 to those with field hockey. Two hundred sixty three of the basketball questionnaires were returned for a response rate of 35.02 percent. Division I returned 108 questionnaires out of 274 for a rate of 39.42 percent. Division II returned 68 questionnaires out of 179 for a rate of 37.99 percent. Division III returned 87 questionnaires out of 298 for a rate of 29.19 percent. For field hockey, 101 questionnaires were returned for a response rate of 40.08 percent. Division I returned 40 questionnaires out of 90 for a rate of 44.44 percent. Division III returned 61 questionnaires out of 162 for a rate of 37.65 percent. Field hockey no longer has a Division II. These data are shown in Table 8 on page 34.

The positions held by each of the respondents are shown in Table 9 on page 35. Only nine, or 2.5 percent were not coaches: eight athletic directors and one national selection committee member. The category, C,O or Coach, Other, includes coaches who were also officials, and/or members of other polls such as those published by USA TODAY and UPI.

Data Analysis

Throughout this research, the data have been evaluated in terms of groups: basketball, total (BT), field hockey, total (HT), and the divisions within each sport. For the purposes of this

Table 8. Questionnaire return rates, by Division.

Basketball

Division	Total	Returned	% Returned
Div 1	274	108	39.42
Div 2	179	68	37.99
Div 3	298	87	29.19
Total	751	263	35.02

Field Hockey

Division	Total	Returned	% Returned
Div 1	90	40	44.44
Div 3	162	61	37.65
Total	252	101	40.08

Table 9. Respondents' positions, 364 total respondents.

No.	Positions
251	Coach (C)
8	Athletic Director (AD)
35	C, AD
16	C, AD, Regional Selection Committee (RSC)
2	C, AD, RSC, National Selection Committee (NSC)
16	C, NSC
10	C, RSC
8	C, RSC, NSC
1	NSC
16	C, Other

discussion, the term *division* is used to refer to Divisions I, II, and III in basketball and/or Divisions I and III in field hockey. The term *groups* will be used to refer to all seven categories, or groupings of data: BT, HT, and the five divisions: B1, B2, B3, H1, and H3.

Regional Bias

To determine the presence of regional bias by division within a sport, the data were placed into computer files, one division per file. These data were further sorted by regional affiliation. The result of this is shown in Appendices E and F.

Table 8 on page 34 shows the reponse for basketball and field hockey by Divisions. The percentage returned for Division 1 was 39.42, Division 2 was 37.99, and Division 3 was 29.19, for an overall return rate for basketball of 35.02. For field hockey, Division 1 returned 44.44 percent, and Division 3 returned 37.46 percent for an overall rate of 40.08 percent. From this, there does not appear to have been bias in terms of Divisions within a sport.

The comparison of regional response rates by division within each sport showed that there were not significant diferences in the returns. This indicates a lack of bias within each sport across regions in terms of the total number of questionnaires made available from each region.

Frequency Analysis

The result of the frequency analysis is found in Appendices I and J. Each question from both questionnaires is ranked by division. An example, from Division I in basketball, is shown in Table 10 on page 37. The frequency - percentage tables that were generated are in Appendices G and H. These data were used to begin to evaluate the selection criteria to be used in the model.

The initial selection was to consider those questions that fell in the upper 50 percent of the distribution, that is, those questions following the logical break in the curve which fell

Table 10. Question numbers and ranks from Division I basketball.

OBS	RANK	DIV	QUES
1	463	B1	Q6
2	452	B1	Q11
3	429	B1	Q9
4	428	B1	Q10
5	425	B1	Q1
6	425	B1	Q5
7	420	B1	Q4
8	410	B1	Q7
9	410	B1	Q12
10	390	B1	Q40
11	387	B1	Q41
12	370	B1	Q24
13	368	B1	Q13
14	367	B1	Q30
15	366	B1	Q20
16	366	B1	Q21
17	366	B1	Q25
18	365	B1	Q31
19	364	B1	Q2
20	363	B1	Q22
21	363	B1	Q23
22	360	B1	Q28
23	353	B1	Q26
24	353	B1	Q27
25	353	B1	Q29
26	353	B1	Q38
27	353	B1	Q39
28	346	B1	Q18
29	346	B1	Q19
30	346	B1	Q36
31	346	B1	Q37
32	342	B1	Q16
33	321	B1	Q14
34	320	B1	Q32
35	318	B1	Q33
36	307	B1	Q3
37	300	B1	Q17
38	295	B1	Q35
39	293	B1	Q8
40	293	B1	Q34
41	180	B1	Q15

closest to 50 percent. The actual cumulative percentages of the midpoints at this break in the curve ranged from a low of 46.88 percent, to a high of 53.66 percent. The number of questions involved ranged from 17 to 24. This was judged to be coarser than necessary; it seemed that many of the questions within this wide range were dissimilar. For contrast, the highest logical break in the curve was examined. The range of actual percentages ran from 93.75 percent to 97.56 percent, with the number of questions ranging from 2 to 6. It appeared that this result was too fine; it was doubted that a reasonable comparison of the model among divisions could be made from two variables. Following this, the logical break in the curve falling closest to 80 percent was evaluated. The ranges for this percentile were: 78.05 to 92.68 percent, and 3 to 10 questions. As the number of questions among the groups appeared to be quite rough using the break in the data, it was decided to use the value from the frequency - percentage tables closest to 80.00 percent as the point from which to evaluate the variables to be included in the model. In addition, the 85th and 90th percentiles were looked at, see Appendices I and J, but the 80th percentile appeared to be the best point at which to begin. Table 11 on page 39 shows the cumulative percentage and the number of questions for each of the percentiles evaluated.

It should be noted that a decision was made to use all of the questions which fell within the upper 20th percent, regardless of the number of groups in which each question occurred. The assumption was made to allow for as much information as possible to be included initially.

Evaluation of Selection Criteria

Evaluating the 80th percentile regardless of logical breaks in the curve appeared to be the point at which the questions were the most similar, across groups. The distribution of the question numbers found in the 80th percentile are listed in Table 12 on page 40. The question numbers are shown in order of rank for each group. The general content of each of these questions is explained in Table 13 on page 41. The kind of information evaluated for ranking

Table 11. Cumulative percentage and number of questions for initial selections.

Break in Curve	GROUPS						
	BT	HT	B1	B2	B3	H1	H3
50 %	51.22[1] 21[2]	50.00 17	48.78 22	53.66 24	51.22 23	46.88 19	53.13 17
80 %	78.05 9	81.25 9	78.05 9	92.68 3	78.05 10	90.63 3	81.25 7
Highest	97.56 2	93.75 4	97.56 2	97.56 3	97.56 3	93.75 3	93.75 6
80 %[3]	80.49 9	81.25 9	78.05 9	85.37 11	80.49 9	81.25 12	81.25 7

[1]Actual cumulative percentage value of midpoint
[2]Number of questions in the percentile
[3]Closest to 80 percent, disregarding logical breaks in the curve

Table 12. Distribution of questions from 80th percentile.

BT[1]	B1	B2	B3	HT	H1	H3
1002[2]	392	257	327	387	157	232
80.49[3]	78.05	85.37	80.49	81.25	81.25	81.25
6[4]	6	6	6	6	6	21
11	11	10	11	11	11	22
9	9	11	9	1	1	6
10	10	1	1	21	10	10
1	1	9	5	22	21	11
5	5	40	10	10	12	1
4	4	7	4	27	22	17
12	7	12	12	2	9	
7	12	4	7	9	27	
		5			5	
		41			2	
					28	

[1]BT = Basketball, all divisions
B1 = Basketball, Division I
B2 = Basketball, Division II
B3 = Basketball, Division III
HT = Field Hockey, all divisions
H1 = Field Hockey, Division I
H2 = Field Hockey, Division III

[2]Interval midpoint, minus 3 points

[3]Actual percentage of interval midpoint minus 3 points

[4]Question number from questionnaire

Table 13. Content of questions from the 80th percentile.

No.[1]	Question Content	Basketball	Field Hockey
1	Comparing records vs common opponents	same	
2	Comparing scores of games vs common opponents	same	
4	Strength of schedule - beginning	same	
5	Strength of schedule - during	same	
6	Win/loss record vs Top 20	same	
7	Credible losses	same	
9	Win/loss record - week	same	
10	Win/loss record - season	same	
11	Win/loss record vs 1-10	same	
12	Win/loss record vs 11-12	same	
17	No equivalent		Win > Tie > Loss
21	No equivalent		Shots on goal - for
22	No equivalent		Shots on goal - against
27	No equivalent		Penalty corners - for
28	No equivalent		Penalty corners - against
40	Field goal % - for	No equivalent	
41	Field goal % - against	No equivalent	

[1]Listed in order of question number from the questionnaires

teams in basketball or field hockey is not sport dependent. In general, both sports use information gathered from Section I in the questionnaire core data, then add a measurement of performance quality, primarily an offensive statistic. The one exception is that in field hockey games may end in a tie while ties cannot occur in basketball. Table 12 on page 40 lists the relationships of the question numbers across groups. The ranked question numbers for the complete study are found in Appendices I and J.

Considering all groups, it can be seen from Table 12 on page 40 that 6 of the 7 rated question 6, *record versus Top 20 teams*, as the most important variable. The seventh group rated question 6 as third most important. Questions numbered 1, 10, and 11 were found in all 7 groups. These are: *comparing records against common opponents*, *season win/loss record* and *win/loss record against the Top 10 teams*, respectively.

The questions described in Table 13 on page 41 became the independent variables for the model. The preliminary model was as shown below.

POS = *f* (record against top 20, record against common opponents, strength of schedule - beginning, strength of schedule - during, credible losses, record - season, record against top 20 teams 1-10, record against top 20 teams 11-20, scoring margin, record - week, and offensive statistic)

Control Data

The data sets used as control data were described in Chapter III. Appendix L contains the actual data sets used to build the control set. The polls created by the evaluators are shown in Table 14 on page 43.

To see if the ranges of the positions assigned by the evaluators may have been significant in some way, the polls were generated a second time using the average position value,

Table 14. Polls generated by evaluators from the control data sets.

Data Set 1			Data Set 2		
POS[1]	TEAM	PTS[2]	POS	TEAM	PTS
1	Derring	177	1	Derring	178
2	Randolph	168	2	Randolph	162
3	Holden	166	3	Whittemore	146
4	Agnew	131	4	Norris	145
5	Norris	130	5	Sandy	141
5	Whittemore	130	6	Cowgill	122
7	Hutcheson	124	7	Holden	121
8	Davidson	106	8	Hutcheson	120
8	Sandy	106	9	Patton	96
10	Cowgill	102	10	Smyth	91
11	Henderson	81	11	Agnew	86
12	McBryde	80	12	Seitz	83
13	Smyth	56	13	Davidson	75
14	Williams	55	14	Price	68
14	Price	55	15	McBryde	58
14	Wallace	55	15	Henderson	58
14	Seitz	55	17	Wallace	54
18	Patton	52	18	Williams	46
19	Pamplin	44	19	Pamplin	25
20	Robeson	36	20	Robeson	14

[1]POS = position
[2]PTS = sum of the point values assigned to position

exclusive of a point assignment. For example, in the second data set Derring was assigned position 1 eight times, and position 3 once. The *average* position, then is 1.2222. Ordering the average position values yielded little change in the polls. In the second poll, positions 6 and 7, Cowgill and Holden, changed spots. In the first poll, where there were three sets of ties, the changes were in the second set. Price, Seitz, and Wallace remained tied at position 14 with average position values of 14.8889, Patton moved to position 17 with a value of 15.2222, and Williams fell to position 18 with a value of 15.8889. The other two sets of ties remained unchanged. These minor changes in the poll positions indicated that the ranges of positions assigned by the evaluators to the teams had little effect on the final polls. As a result of this, and the fact that selection committees use the points system, the decision was made to use the point method for the study.

The correlation matrix was created for the second control data set to examine the relationships between the variables. The data indicated a very strong relationship between two sets of variables. The strongest is between strength of schedule, beginning of season (SOSB) and strength of schedule, during season (SOSD). This is to be expected; if the method used to predict the preseason strengths is reliable, it would be expected that, barring some unforeseen catastrophe to several teams, the schedule ratings would not fluctuate to a great degree during the season. Given the strong relationship between these variables, and the fact that for polling purposes, information *during* the season is more relevant than *preseason* data, the decision was made to drop SOSB from the model.

The other highly correlated variables were season win/loss record (RSN) and scoring margin (SCO). This is reasonable given the situation. When considering teams for a top twenty poll, it would be expected that the RSN would be very good, that is, significantly more wins than losses. With SCO being the difference between points scored and points allowed, to win a team must have a positive SCO. With many wins versus few losses, the SCO would tend to get large. It is expected that, on the average, the top twenty teams would not lose many games by large margins, which would indicate a greater positive SCO. Although the information projected by each variable is related, they provide different indices toward overall

performance. For example, if several teams had the same RSN, all other things being equal, SCO margin may provide the information which would enable a tie to be broken. Due to this, both variables were kept in the model.

A strongly negative relationship was found between credible losses (CL) and record versus teams in the top twenty poll (R20). It would be expected that with few losses there would be less chance that many would be "credible."

An obvious, additive relationship existed between R20, and record versus teams ranked 1 through 10 (R10) and those ranked 11 through 20 (R11). Since more information is provided by the split variables, R10 and R11, it was decided to drop R20 from the model.

A stepwise regression was run on the second control data set using the points (PTS) from the evaluators' poll as the dependent variable. Since the model is not to be used for prediction, regression analysis was not necessary. However, it was suspected that each variable within the model was not weighted equally as selection criteria by the evaluators. As such, it was necessary to attempt to understand the relationship of the independent variables to the dependent variable. The premise was that the regression would indicate the more "important" variables, which would aid in the weighting process.

Fifteen variations on the model were run against the second control set, using different combinations of the variables. This was done to determine if some combinations of variables were better than others to "predict" the same poll as the evaluators. The equations, along with the polls may be found in Appendix M. Where variable substitution was done on collinear or additive variables, the equations yielded the same ordering of teams, but with different point values. However, the order of none of the equations was equal to that given by the evaluators. From this can be inferred that the evaluators did not use all of the information made available to them and/or did not weight each of the variables equally.

The regression equation was ultimately rejected for three reasons: (a) it was composed of only four variables, two of which were possibly collinear, (b) the model is not to be used for prediction, and (c) the objective was to provide standard baseline data to the selection committees. Using the model with unweighted variables produced a poll differing greatly from the

one developed from the evaluators. Some of this variation can be explained as the difference in the individual weights assigned by each evaluator, some by the fact that some of the variables appear to require weights by definition. For example if the record against the top 10 teams is the same weight as the record against teams 11-20, then of what point is the differentiation? To attempt to reduce the amount of pure subjectivity introduced in assigning weights, a portion of Giles' objectives-weighting procedure was applied, with the variables approximating objectives in the matrix.

The Model

Following the decisions discussed above, the model was:

$$POS = RCO - SOSD - CL + RSN + R10 + R11 + SCO + STAT + RWK$$

with the order of the variables being not significant. The first step in developing weights for each variable was to determine the relative importance assigned by the questionnaire respondents. A matrix was designed with the divisions from the sports as the rows and the variables as the columns. This is shown in Table 15 on page 47. The total point values assigned each variable from the questionnaire data was divided by the total number of questionnaires returned within the division. The average score for each variable is the value in each cell of the matrix. Table 16 on page 48 shows the average score for each cell, rounded to one significant digit. Each column was then summed and divided by five, the total number of divisions between the two sports. This value, rounded to two significant digits, is the weight for each variable.

Since the literature suggested that weights are most meaningful when occurring within a range of between 0 and ± 1, it was decided to scale the weights. The largest weight, 4.22,

Table 15. Weighting matrix showing average calculations.

DIVISION[1]	RCO	SOSD	CL	RSN	R10	R11	SCO	STAT[2]	RWK
B1 (108)	425[3]	425	410	428	452	410	364	389	429
B2 (68)	268	261	262	280	278	262	242	261	265
B3 (87)	361	348	329	347	367	335	319	318	362
H1 (40)	177	160	154	165	183	163	159	163	161
H3 (61)	244	217	225	246	245	222	230	254	227

[1]Division with the number of responses for that division.
[2]STAT is a combination of two questions: for and against. The difference between the two was halved, then added to the lower value of the two to get the cell value.
[3]The total number of points for each variable, from the questionnaire data.

Table 16. Weighting matrix showing average values per cell.

	RCO	SOSD	CL	RSN	R10	R11	SCO	STAT	RWK
B1	3.9	3.9	3.8	4.0	4.2	3.8	3.4	3.6	4.0
B2	3.9	3.8	3.9	4.1	4.1	3.9	3.6	3.8	3.9
B3	4.1	4.0	3.8	4.0	4.2	3.9	3.7	3.7	4.2
H1	4.4	4.0	3.9	4.1	4.6	4.1	4.0	4.1	4.0
H3	4.0	3.6	3.7	4.0	4.0	3.6	3.8	4.2	3.7
÷ 5	4.06	3.86	3.82	4.04	4.22	3.86	3.70	3.88	3.96

was set to 1, and all others scaled accordingly, using the ratio $x \div 4.22$, where x is the unscaled weight for each variable. The final model, then, became:

$$POS = R10 + 0.962(RCO) - 0.915(SOSD) - 0.905(CL) + 0.957(RSN) + 0.915(R11) + 0.877(SCO) + 0.919(STAT) + 0.938(RWK)$$

The derivations of each variable is explained in Table 17 on page 50.

The algorithm was run using the two control data sets. The results were not exactly the same as either of the polls created by the evaluators. The polls using the model are shown in Table 18 on page 51. It is believed that the majority of the discrepancy can be explained by the fact that each individual selection method utilized by the evaluators cannot be accounted for in the model. What is provided is a standard set of data, weighted by average values suggested by the questionnaire data, which will yield an ordering of some specified number of teams, greater than 20. This "poll" would then be examined by a selection committee using whatever subjective means deemed necessary to arrive at a consensus top twenty poll.

Table 17. Variable derivations.

R10: Win (W)[1]/Loss (L) record vs top 20 teams 1-10; expressed as a proportion: $W \div (W + L)$

RCO: W/L vs common opponents; expressed as a proportion: $W \div (W + L)$

SOSD: Strength of schedule, during the season; expressed as a proportion, $X \div Y$, where X is the strength value, and Y is the total number of possible values.

CL: Credible losses; number of losses in basketball, or losses and/or ties in field hockey, to a higher ranked team by < 5 points in basketball, or ≤ 2 points in field hockey; expressed as an integer.

RSN: W/L for season; expressed as a proportion: $W \div (W + L)$

R11: W/L for top 20 teams 11-20; expressed as a proportion: $W \div (W + L)$

RWK: W/L for current week; expressed as a proportion: $W \div (W + L)$

SCO: Scoring margin; expressed as a proportion: $SCO \div 100$

STAT: For basketball: field goal percentage for, minus field goal percentage against; for field hockey, scoring percentage (S%) where S% = goals made (GM) ÷ shots on goal (SOG)

[1]The derivations shown here are developed for basketball. For field hockey where ties (T) may occur, all winning "percentages" may be calculated as: $(W + .5T) \div (W + T + L)$. While this is not truly a winning percentage, it it is a relative measure whereby the following condition is met: $W > T > L$.

Table 18. Polls generated by the weighted algorithm.

	Data Set 1			Data Set 2	
POS	TEAM	PTS	POS	TEAM	PTS
1	Derring	4.94988	1	Derring	3.6222
2	Cowgill	4.17308	2	Whittemore	3.5993
3	Randolph	3.71425	3	Hutcheson	3.3218
4	Sandy	3.60404	4	Cowgill	3.2901
5	Whittemore	3.56174	5	Sandy	3.2517
6	Williams	3.38181	6	Randolph	2.5925
7	Price	3.25398	7	Price	2.1547
8	Hutcheson	2.96066	8	Seitz	1.4647
9	Agnew	2.60597	9	Patton	1.1440
10	Norris	2.43572	10	Pamplin	1.0871
11	Patton	2.40577	11	Norris	0.9637
12	Smyth	1.92183	12	Smyth	0.7131
13	Pamplin	1.91826	13	Williams	-0.1358
14	Davidson	1.49582	14	Agnew	-0.3264
15	Henderson	1.38425	15	Robeson	-1.4716
16	Robeson	1.35911	16	Henderson	-1.9959
17	McBryde	1.29164	17	McBryde	-2.2406
18	Holden	1.14482	18	Wallace	-3.2109
19	Wallace	1.11460	19	Davidson	-3.8579
20	Seitz	0.91944	20	Holden	-5.1256

Chapter V

Summary, Conclusions, and Recommendations

Summary

The objective of this research was to determine whether a method could be developed to establish continuity in ranking techniques used within women's team sports. The methodologies used by the individuals serving on selection committees were examined, the direct result of which was the development of questionnaires reflecting these methods.

The questionnaires were then used to determine whether the selection methodologies were sport-specific, or division-specific. From the data within the 80th percentile, there does not appear to be any difference between the variables used to generate top twenty polls for basketball and field hockey, or for within the divisions within each sport. Win/loss records, scoring margins, credible losses, and strength of schedules were rated most important, along with a measure of performance quality. The performance quality indicator was taken as a statistic kept in terms of for and against; for example, in basketball, field goal percentage for, field goal percentage against, or allowed. This index, then, was used as a relative measure of overall team performance quality.

For field hockey, a statistic related to that used in basketball was developed based on goals scored and shots on goal. It is believed that a similar statistic can be developed for each of the other women's team sports, making the model adaptable and applicable for all NCAA women's team sports.

The model is not a prediction tool, per se. What the model is is a mathematical equation which evaluates some specific number of teams from a standard set of selection criteria. These teams are ordered based on a series of weighted variables generating a poll derived from quantitative data.

Conclusions

Selection committees are provided with a relative measure of comparison; each team is compared against every other team based on the same standardized information. The number of teams to be compared is limited only by the number of teams in the data set from which the model operates. By generating a poll with some number greater than 20 teams, the selection committees could use whatever subjective means deemed necessary to arrive at a concensus top twenty poll.

This method significantly reduces potential bias. The simplest means is by providing for comparisons regardless of regional affiliation. The model also reduces bias due to the standardization of these comparisons. From the set of variables used by the model, nothing may be omitted or disregarded from team to team, or region to region.

The model cannot take into account all of the individual selection methodologies. The weights assigned to the variables are an attempt to reconcile this by the fact that they are averages based on the study data, reflecting the methods of 364 coaches and selection committee members associated with basketball and field hockey. The weights could have been "arbitrarily" assigned, and could yet be changed, but it is believed that the method used pro-

vides the most objective means for reducing the amount of pure subjectivity introduced into the model.

The limitations of the model will be in how much an operator is willing to input into the data set. Ideally, entering the data for 50-60 teams would provide the selection committees with a greater number of comparisons, potentially yielding more meaningful polls.

The equation tends to provide a greater delineation of poll positions when there have been more games played. This would appear to indicate that earlier in a season there would be more ties in the ordering of the teams due to very similar records. While individuals generally do not assign ties to teams while ranking, ties do occur in the polls due to the range of positions one team may occupy in the various individual polls. This is indicated in the small study example: the first control data set had three sets of ties, and the second had one. It is believed that, more than anything, this demonstrates the difficulty in an objective ranking based on identical, or very similar data. This is not judged to be a problem in that the ties would still be relative in terms of overall position, and the selection committee would break them based on subjective criteria.

Strength of schedule must be determined by some means external to the model. It is recommended that the NCAA's schedule power ranking program be used. Not only is this program somewhat of a standard for use within the NCAA, but in reviewing the literature for this study, no other means for arriving at relative strength of schedule not requiring extensive networking techniques was discovered. Inasmuch as the NCAA's program already exists and is used, it was judged not necessary to duplicate that effort.

Recommendations for Future Study

It is recommended that:

1. The study be replicated within 10 years to determine whether the selection criteria has changed,

2. An interactive computer program be written to facilitate use of the model,
3. A comparison be made between the responses from the coaches and those from the selection committee members to determine if there are perceived differences in the methods, and
4. A future study consider game site as a variable.

References

Abraham, B. & Ledolter, J. (1983). *Statistical methods for forecasting*. New York: John Wiley and Sons.

Aitchison, J. & Dunsmore, I.R. (1975). *Statistical prediction analysis* Cambridge: Cambridge University Press.

Bellman, R. (1977). Dynamic programming and Markovian decision processes, with application to baseball. In S.P. Ladany & R.E. Machol (Eds.), *Optimal strategies in sports* (pp. 77-85). New York: North Holland.

Bennett, J.M. & Flueck, J.A. (1983). An evaluation of major league baseball offensive performance models. *The American Statistician*, *3*(1), 76-82.

Dillman, D.A. (1978). *Mail and telephone surveys*. New York: John Wiley and Sons.

Dorf, R.C. (1977). *Introduction to computers and computer science*. San Francisco: Boyd and Fraser.

Evans, J.R., Hebert, J.E., & Deckro, R.F. (1984). Play ball-The scheduling of sports officials. *Perspectives in Computing*, *4*(1), 18-29.

Friedman, A. (1978). *The world of sports statistics*. New York: Atheneum.

Giles, R.H., Jr. (1986). Extensions of the maximization decision strategy for wildlife managers. Unpublished manuscript, Virginia Polytechnic Institute and State University, Department of Fisheries and Wildlife Sciences, Blacksburg, VA.

Gupta, S.S. & Nagel, K. (1971). On some contribution to multiple decision theory. In S.S. Gupta, & J. Yackel (Eds.). *Statistical Decision Theory and Related Topics*. Proceedings of a symposium held at Purdue University Nov. 23-25, 1970.

Haberman, S.J. (1974). *The analysis of frequency data*. Chicago: The University of Chicago Press.

Jenkins, G.M. (1976). Practical experience with modelling and forecasting time series. In O.D. Anderson (Ed.) *Forecasting*. Proceedings of the Institute of Statisticians Annual Conference, Cambridge 1976.

Kerlinger, F.N. (1973). *Foundations of behavioral research.* New York: Holt, Rinehart and Winston.

Leake, R.J. (1976). A method for ranking teams: With an application to college football. In R.E. Machol, S.P. Ladany, & D.G. Morrison (Eds.)., *Management Science in Sports* (pp. 27-46). New York: North Holland.

Machol, R.E. (1977). A Word to the Non-Mathematician. In S.P. Ladany & R.E. Machol (Eds.), *Optimal strategies in sports* (pp. xvi-xx). New York: North Holland.

McDonald, C. (1971). On approximating constants required to implement a selection procedure based on ranks In S.S. Gupta, & J. Yackel (Eds.). *Statistical decision theory and related topics.* Proceedings of a symposium held at Purdue University Nov. 23-25, 1970.

Mitchell, C.R. & Michel, A.J. (1977). The valuation of a baseball player. In S.P. Ladany & R.E. Machol (Eds.), *Optimal strategies in sports* (pp. 42-45). New York: North Holland.

Mosteller, F. & Tukey, J.W. (1977). *Data analysis and regression.* Reading, MA.: Addison-Wesley.

National Collegiate Athletic Association. (1979). *Manual of the National Collegiate Athletic Association.* Shawnee Mission, KS: Author.

National Collegiate Athletic Association. (1985). *National collegiate championships handbook - women's basketball* Mission, KS: Author.

National Collegiate Athletic Association. (1983). *National collegiate championships handbook - field hockey.* Mission, KS: Author.

National Collegiate Athletic Association. (1986). *NCAA NEWS.* March 12, 1986. Mission, KS: Author.

Ott, L. (1977). *Introduction to statistical methods and data analysis.* North Scituate, MA.: Duxbury Press.

Price, B. & Rao, A.G. (1977). A model for evaluating player performance in professional basketball. In S.P. Ladany & R.E. Machol (Eds.), *Optimal strategies in sports* (pp. 116-122). New York: North Holland.

Rosner, B. (1976). An analysis of professional football scores. In R.E. Machol, S.P. Ladany, & D.G. Morrison (Eds.). *Management Science in Sports* (pp. 59-78). New York: North Holland.

Ryan, F., Francia, A.J., & Strawser, R.H. (1977). Informational analysis and professional football. In S.P. Ladany & R.E. Machol (Eds.), *Optimal strategies in sports* (pp. 112-115). New York: North Holland.

SAS Institute, Inc. (1983). *SAS companion for the VM/CMS operating system, 1983 edition.* Cary, NC.: SAS Institute.

SAS Institute, Inc. (1985). *SAS user's guide: statistics, Version 5 Edition.* Cary, NC.: SAS Institute.

Sawusch, M.R. & Summers, T.A. (1984). *1001 things to do with your IBM PC.* Blue Ridge Summit, PA: Tab Books.

Sirgy, J.M. & Giles, R.H., Jr. (1985, March). *A general system model of goal formulation in human activity systems.* Paper presented at the meeting of the American Association for the Advancement of Science, San Francisco, CA.

Smith, J.H. (1977). Adjusting baseball standings for strength of teams played. In S.P. Ladany & R.E. Machol (Eds.), *Optimal strategies in sports* (pp. 48-49). New York: North Holland.

Ushakov, I.A. (1976). The problem of choosing the preferred element: An application to sport games. In R.E. Machol, S.P. Ladany, & D.G. Morrison (Eds.), *Management Science in Sports* (pp. 153-161). New York: North Holland.

Van Valkenburg, J.M. (1986, January 22). Top teams invariably have the toughest schedules. *NCAA News*, p. 3.

Webster's new collegiate dictionary. (1981). Springfield, Mass.: G. and C. Merriam.

Whittle, P. (1983). *Prediction and regulation by linear least squares methods.* Minneapolis: University of Minnesota Press.

Appendix A. Glossary

a priori: Based on theory instead of experience or experiment, or before examination or analysis (Webster's New Collegiate Dictionary, 1981).

Bias: A systematic error introduced into sampling by encouraging one outcome over others (Webster's New Collegiate Dictionary, 1981).

Common Opponents: Opponents shared by more than one team. For example: Team A and Team B do not play each other. However, they both play Team C. Team C is then a common opponent to both Team A and Team B.

Credible losses: Losing to a team of higher ranking, but by a much less margin than expected. For basketball, credible was defined as a difference of 5 points. For field hockey, the difference is 2 points.

Index: A number that is a composite of two or more other numbers (Kerlinger, 1973) expressed as a ratio or proportion. A percentage is a variation of a proportion.

Individual Bias: Bias introduced by an individual selection committee member not generally associated to that of others within a region or division.

Intransitivity: *Transitivity* is related to the mathematical equality: if A = B, and B = C, then C = A. Intransitivity occurs, then, when A beats B, and B beats C, but C beats A.

Manual File: A collection of related data treated as a unit, not residing on a computer.

Model: A qualitative or quantitative representation of a process. A model is used to simulate or predict based on a given set of factors, or variables (Dorf, 1977). In terms of this study, the model will be used to simulate top twenty polls.

Multicollinearity: A condition which occurs in regression when independent variables are highly correlated, making the correlation matrix nearly linear. This is an indication that more than one variable is predicting the same thing, linearity, and the effect of each cannot be accurately determined. For this study, collinearity was used to justify dropping variables from the model which were nearly linear or additive.

National Selection Committee: Committee formed to choose those teams to appear as members of national weekly ranking polls. Also chooses the teams to appear in the national championships.

Population: A set of subjects with an identifiable common characteristic. In this study, the population was all of the coaches and selection committee members for women's basketball and field hockey affiliated with the National Collegiate Athletic Association.

Ranking: A rated position relative to other teams participating in the same sport. In terms of this study, teams are ranked from 1 to 20, with 1 being the highest and 20 being the lowest.

Regional Bias: Bias introduced regionally, generally through information which is weighted toward teams in a specific region. The study could have been biased regionally by a statistical difference in the number of questionnaires returned from one or more regions in a division.

Regional Selection Committee: Committee formed to choose those teams to to be considered by the national selection committee for the national weekly ranking polls.

Sample: A subset of a population. The sample for this study is the returned questionnaires.

Scoring Margin: The difference obtained by subtracting the lower score from the the higher score in one game, summed for all games.

Set: Generally, a grouping of data. In terms of this study, the data were grouped by divisions within a sport. Same as Data Set.

Strength-of-schedule: Generally, considering how many "strong" teams are scheduled during one team's season. "Strong" would be teams such as those in the "top twenty" polls, or top conference teams. While a team's schedule does not change during the course of a season, "strength-of-schedule" may as opponents move in and out of the "top twenty."

Top Twenty Poll: A list denoting a ranking of women's teams for a sport. In terms of this study, the ranking is of 20 teams, determined by a polling of selection committees, thus "Top Twenty Poll."

Women's Team Sports: The NCAA recognizes six sports as team sports for women: basketball, field hockey, lacrosse, soccer, softball, and volleyball.

Appendix B. Basketball Questionnaire

Determining "Top Twenty" Polls

The following questionnaire is an attempt to determine the criteria used by ranking committees while developing weekly "Top Twenty" polls for their sports. Following the instructions for each section, please answer each question based on your individual selection criteria.

Please return to completed questionnaire by July 10, 1985 to:

D.A. Rasberry
c/o Dr. M.L. Driscoll
War Memorial Gym
Virginia Tech
Blacksburg, VA 24061

Definitions:

1. Common opponents: Team A and Team B do not play each other. However, they both play Team C. Team C is then a common opponent to both Team A and Team B.

2. Scoring margin: The difference obtained by subtracting the lower score from the higher score in one game.

3. Strength-of-schedule: Generally, considering how many "strong" teams (i.e., "top twenty" teams, or top conference teams) are played during one team's schedule. While a team's schedule does not change during the course of a season, "strength-of-schedule" may as opponents move in and out of the "top twenty."

4. Credible wins/losses: Generally, losing to a team but by a much less margin than expected. For example, losing to a top-ranked team by a score of 88 to 87. Subjective term.

I. This section is addressing relative importance of a particular factor. You choose whether each item in your consideration for the poll is **not considered** (NC), **not important** (NI), is **somewhat important** (SI), **important** (I), or is **very important** (VI). Circle the number that corresponds with that choice. Please circle only one choice for each question.

How important is: (choose only one)	NC	NI	SI	I	VI
1. Comparing records against common opponents?	1	2	3	4	5
2. Comparing scores of games against common opponents?	1	2	3	4	5
3. Scoring margins? (e.g. is 85-70 as good or bad as 50-35?)	1	2	3	4	5
4. Strength-of-schedule, beginning of season?	1	2	3	4	5
5. Strength-of-schedule as opponents on that schedule move in or out of the "top twenty?" (Does strength-of-schedule change during the season?)	1	2	3	4	5
6. Games won or lost against "top twenty" opponents?	1	2	3	4	5
7. Credible losses?	1	2	3	4	5
8. Credible wins?	1	2	3	4	5
9. Considering only the current week's win/loss record, or that of the current and the immediately previous week?	1	2	3	4	5
10. Considering the complete win/loss record up to the date of consideration?	1	2	3	4	5
11. Wins/losses against "top ten" (1-10) teams?	1	2	3	4	5
12. Wins/losses against "bottom ten" (11-20) teams of the "top twenty?"	1	2	3	4	5
13. The record if one team plays another more than once?	1	2	3	4	5
14. Subjective comparison of teams from other regions with no common opponents?	1	2	3	4	5
15. Considering first half scores versus complete game scores?	1	2	3	4	5

II. The following two statements may be considered true (yes) or false (no). To the right of each statement please circle the answer that best applies to your consideration of "top twenty" polls.

Statement is True or False: (choose only one)		True	False
16.	A win against a team that is unranked then later is ranked counts the same as a win over a currently ranked team.	1	2
17.	A win/loss against the "number one" ranked team in the first week of the season counts the same as a win/loss against the "number one" team in week ten.	1	2

III. This section concerns the statistics that are kept on the official score sheets. As in Section I, the questions are rated **not considered** (NC), **not important** (NI), **somewhat important** (SI), is **important** (I), or **very important** (VI). Circle the number that corresponds with your choice.

How important is: (choose only one)		NC	NI	SI	I	VI
18.	Field goal attempts - for	1	2	3	4	5
19.	Field goal attempts - against	1	2	3	4	5
20.	Field goals made - for	1	2	3	4	5
21.	Field goals made - against	1	2	3	4	5
22.	Free throw attempts - for	1	2	3	4	5
23.	Free throw attempts - against	1	2	3	4	5
24.	Free throws made - for	1	2	3	4	5
25.	Free throws made - against	1	2	3	4	5
26.	Offensive rebounds - for	1	2	3	4	5
27.	Offensive rebounds - against	1	2	3	4	5
28.	Defensive rebounds - for	1	2	3	4	5
29.	Defensive rebounds - against	1	2	3	4	5
30.	Turnovers - for	1	2	3	4	5
31.	Turnovers - against	1	2	3	4	5
32.	Shots taken from inside the key	1	2	3	4	5
33.	Shots allowed from inside the key	1	2	3	4	5
34.	Shots taken from outside the key	1	2	3	4	5

35.	Shots allowed from outside the key	1	2	3	4	5
36.	Field goal percentage, 1st half - for	1	2	3	4	5
37.	Field goal percentage, 1st half - against	1	2	3	4	5
38.	Field goal percentage, 2nd half - for	1	2	3	4	5
39.	Field goal percentage, 2nd half - against	1	2	3	4	5
40.	Field goal percentage, game - for	1	2	3	4	5
41.	Field goal percentage, game - against	1	2	3	4	5

IV. Circle any of the following positions that apply to you:

1. Coach (of sport concerning this questionnaire)
2. Athletic Director
3. Regional polling committee member
4. National polling committee member
5. Other (Specify below)

Thank you very much for your participation with this questionnaire.

Appendix C. Field Hockey Questionnaire

Determining "Top Twenty" Polls

The following questionnaire is an attempt to determine the criteria used by ranking committees while developing weekly "Top Twenty" polls for their sports. Following the instructions for each section, please answer each question based on your individual selection criteria.

Please return to completed questionnaire by July 10, 1985 to:

D.A. Rasberry
c/o Dr. M.L. Driscoll
War Memorial Gym
Virginia Tech
Blacksburg, VA 24061

Definitions:

1. Common opponents: Team A and Team B do not play each other. However, they both play Team C. Team C is then a common opponent to both Team A and Team B.

2. Scoring margin: The difference obtained by subtracting the lower score from the higher score in one game.

3. Strength-of-schedule: Generally, considering how many "strong" teams (i.e., "top twenty" teams, or top conference teams) are played during one team's schedule. While a team's schedule does not change during the course of a season, "strength-of-schedule" may as opponents move in and out of the "top twenty."

4. Credible wins/losses: Generally, losing to a team but by a much less margin than expected. For example, losing to a top-ranked team by a score of 1 to 0. Subjective term.

I. This section is addressing relative importance of a particular factor. You choose whether each item in your consideration for the poll is **not considered** (NC),**not important** (NI), is **somewhat important** (SI), **important** (I), or is **very important** (VI). Circle the number that corresponds with that choice. Please circle only one choice for each question.

How important is: (choose only one)	NC	NI	SI	I	VI
1. Comparing records against common opponents?	1	2	3	4	5
2. Comparing scores of games against common opponents?	1	2	3	4	5
3. Scoring margins? (e.g. is 15-10 as good or bad as 5-0?)	1	2	3	4	5
4. Strength-of-schedule, beginning of season?	1	2	3	4	5
5. Strength-of-schedule as opponents on that schedule move in or out of the "top twenty?" (Does strength-of-schedule change during the season?)	1	2	3	4	5
6. Games won or lost against "top twenty" opponents?	1	2	3	4	5
7. Credible losses?	1	2	3	4	5
8. Credible wins?	1	2	3	4	5
9. Considering only the current week's win/loss record, or that of the current and the immediately previous week?	1	2	3	4	5
10. Considering the complete win/loss record up to the date of consideration?	1	2	3	4	5
11. Wins/losses against "top ten" (1-10) teams?	1	2	3	4	5
12. Wins/losses against "bottom ten" (11-20) teams of the "top twenty?"	1	2	3	4	5
13. The record if one team plays another more than once?	1	2	3	4	5
14. Subjective comparison of teams from other regions with no common opponents?	1	2	3	4	5
15. Considering first half scores versus complete game scores?	1	2	3	4	5

II. The following five statements may be considered true (yes) or false (no). To the right of each statement please circle the answer that best applies to your consideration of "top twenty" polls.

Statement is True or False: (choose only one)	True	False
16. Field conditions are considered for games where weather may have been a factor in the outcome of the game.	1	2
17. Ties are considered better than a loss but not as good as a win.	1	2
18. Ties are considered the same as a loss.	1	2
19. A win against a team that is unranked then later is ranked counts the same as a win over a currently ranked team.	1	2
20. A win/loss against the "number one" ranked team in the first week of the season counts the same as a win/loss against the "number one" team in week ten.	1	2

III. This section concerns the statistics that are kept on the official score sheets. As in Section I, the questions are rated **not considered** (NC), **not important** (NI), **somewhat important** (SI), is **important** (I), or **very important** (VI). Circle the number that corresponds with your choice.

How important is: (choose only one)	NC	NI	SI	I	VI
21. Shots on goal - for	1	2	3	4	5
22. Shots on goal - against	1	2	3	4	5
23. Goalkeeper saves - for	1	2	3	4	5
24. Goalkeeper saves - against	1	2	3	4	5
25. Defensive saves - for	1	2	3	4	5
26. Defensive saves - against	1	2	3	4	5
27. Penalty corners - for	1	2	3	4	5
28. Penalty corners - against	1	2	3	4	5
29. Long corners - for	1	2	3	4	5
30. Long corners - against	1	2	3	4	5
31. Penalty strokes - for	1	2	3	4	5
32. Penalty strokes - against	1	2	3	4	5

IV. Circle any of the following positions that apply to you:

1. Coach (of sport concerning this questionnaire)
2. Athletic Director
3. Regional polling committee member
4. National polling committee member
5. Other (Specify below)

Thank you very much for your participation with this questionnaire.

Appendix D. Cover Letters to Coaches and Selection Committee Members

College of Education
Division of HPER

Dear Coach,

The National Collegiate Athletic Association recognizes 6 sports as team sports for women: basketball, field hockey, lacrosse, soccer, softball, and volleyball. Selection committees in each sport consult via telephone conference calls to develop weekly rankings, or "Top Twenty Polls." Currently, there is no means for eliminating regional and/or individual bias in these rankings, beyond the selection of competent people for the committees.

As part of a research project, a ranking model is being developed to provide standardized "baseline" data to the selection committees. The study sports are women's basketball and field hockey. The head coaches of institutions participating in at least one of these sports, as well as selection committee members, are being requested to participate in this study by responding to the enclosed questionnaire which has been developed to determine the criteria used while ranking women's athletic teams.

Complete confidentiality will be maintained. Each questionnaire has an identification number for the purpose of checking your name off of a mailing list when your questionnaire is returned. The information gathered from the questionnaire will be used for statistical analysis only.

The results of the questionnaire will be used to determine the criteria that are used in ranking women's team sports. The model, then, will serve to provide standard data to selection committees for further discussion in the ranking process. This will eliminate much of the potential bias.

Please do not hesitate to write or call if there are questions. The telephone number is (703) 961-6561.

Thank you very much for your assistance.

D. Ann Rasberry

Margaret L. Driscoll Professor
War Memorial Gym
Virginia Tech
Blacksburg, VA 24061

College of Education
Division of HPER

Dear Selection Committee Member,

The National Collegiate Athletic Association recognizes 6 sports as team sports for women: basketball, field hockey, lacrosse, soccer, softball, and volleyball. Selection committees in each sport consult via telephone conference calls to develop weekly rankings, or "Top Twenty Polls." Currently, there is no means for eliminating regional and/or individual bias in these rankings, beyond the selection of competent people for the committees.

As part of a research project, a ranking model is being developed to provide standardized "baseline" data to the selection committees. The study sports are women's basketball and field hockey. The head coaches of institutions participating in at least one of these sports, as well as selection committee members, are being requested to participate in this study by responding to the enclosed questionnaire which has been developed to determine the criteria used while ranking women's athletic teams.

Complete confidentiality will be maintained. Each questionnaire has an identification number for the purpose of checking your name off of a mailing list when your questionnaire is returned. The information gathered from the questionnaire will be used for statistical analysis only.

The results of the questionnaire will be used to determine the criteria that are used in ranking women's team sports. The model, then, will serve to provide standard data to selection committees for further discussion in the ranking process. This will eliminate much of the potential bias.

Please do not hesitate to write or call if there are questions. The telephone number is (703) 961-6561.

Thank you very much for your assistance.

D. Ann Rasberry

Margaret L. Driscoll Professor
War Memorial Gym
Virginia Tech
Blacksburg, VA 24061

Appendix E. Basketball Questionnaire Response

Basketball - All Divisions

Region [1]	Total	Returned	% Returned
Div 1	274	108	39.42
Div 2	179	68	37.99
Div 3	298	87	29.19
Total	751	263	35.02

[1]Based on NCAA Sports Sponsorship 1985-86 Winter Sports Supplement to the National Collegiate Championships Handbook. Region totals may include committee representatives in addition to school representatives.

Basketball - Division 1

Region [1]	Total	Returned	% Returned
East	97	36	37.11
Mideast	72	28	38.89
Midwest	59	24	40.68
West	46	20	43.48
Totals	274	108	39.42

[1]Based on NCAA Sports Sponsorship 1985-86 Winter Sports Supplement to the National Collegiate Championships Handbook. Region totals may include committee representatives in addition to school representatives.

Basketball - Division 2

Region [1]	Total	Returned	% Returned
New England	18	5	27.77
East	33	14	42.42
So. Atlantic	24	9	37.50
South	25	9	36.00
Great Lakes	24	8	33.33
No. Central	13	5	38.46
So. Central	24	12	50.00
West	18	6	33.33
Totals	179	68	37.99

[1]Based on NCAA Sports Sponsorship 1985-86 Winter Sports Supplement to the National Collegiate Championships Handbook

Basketball - Division 3

Region [1]	Total	Returned	% Returned
Northeast	56	14	25.00
East	42	12	28.57
Mid-Atlantic	32	9	28.13
Atlantic	36	11	30.56
South	38	10	26.32
Great Lakes	25	10	40.00
Central	36	9	25.00
West	33	12	36.37
Totals	298	87	29.19

[1]Based on NCAA Sports Sponsorship 1985-86 Winter Sports Supplement to the National Collegiate Championships Handbook

Appendix F. Field Hockey Questionnaire Response

Field Hockey - All Divisions

Region [1]	Total	Returned	% Returned
Div 1	90	40	44.44
Div 3	162	61	37.65
Total	252	101	40.08

[1]Based on NCAA Sports Sponsorship 1983 Fall Sports Supplement to the National Collegiate Championships Handbook. Region totals may include committee representatives in addition to school representatives.

Field Hockey - Division 1

Region [1]	Total	Returned	% Returned
New England	16	7	43.75
Mideast	19	7	36.84
Mid-Atlantic	11	5	45.45
South	18	8	44.44
Midwest	21	11	52.38
West	5	2	40.00
Totals	90	40	44.44

[1]Based on NCAA Sports Sponsorship 1983 Fall Sports Supplement to the National Collegiate Championships Handbook. Region totals may include committee representatives in addition to school representatives.

Field Hockey - Division 3

Region [1]	Total	Returned	% Returned
Northeast	47	16	34.04
Mid-Atlantic	23	8	34.78
Pennsylvania	26	10	38.46
South	24	10	41.67
Great Lakes	42	17	40.48
Totals	162	61	37.65

[1]Based on NCAA Sports Sponsorship 1983 Fall Sports Supplement to the National Collegiate Championships Handbook. Region totals may include committee representatives in addition to school representatives.

Appendix G. Frequency - Percentage Table for Basketball

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
		Basketball, All Divisions		
0	0	0	0.00	0.00
5	0	0	0.00	0.00
10	0	0	0.00	0.00
15	0	0	0.00	0.00
20	0	0	0.00	0.00
25	0	0	0.00	0.00
30	0	0	0.00	0.00
35	0	0	0.00	0.00
40	0	0	0.00	0.00
45	0	0	0.00	0.00
50	0	0	0.00	0.00
55	0	0	0.00	0.00
60	0	0	0.00	0.00
65	0	0	0.00	0.00
70	0	0	0.00	0.00
75	0	0	0.00	0.00
80	0	0	0.00	0.00
85	0	0	0.00	0.00
90	0	0	0.00	0.00
95	0	0	0.00	0.00
100	0	0	0.00	0.00
105	0	0	0.00	0.00
110	0	0	0.00	0.00
115	0	0	0.00	0.00
120	0	0	0.00	0.00
125	0	0	0.00	0.00
130	0	0	0.00	0.00
135	0	0	0.00	0.00
140	0	0	0.00	0.00
145	0	0	0.00	0.00
150	0	0	0.00	0.00
155	0	0	0.00	0.00
160	0	0	0.00	0.00
165	0	0	0.00	0.00
170	0	0	0.00	0.00
175	0	0	0.00	0.00
180	0	0	0.00	0.00
185	0	0	0.00	0.00
190	0	0	0.00	0.00
195	0	0	0.00	0.00
200	0	0	0.00	0.00
205	0	0	0.00	0.00
210	0	0	0.00	0.00
215	0	0	0.00	0.00
220	0	0	0.00	0.00
225	0	0	0.00	0.00
230	0	0	0.00	0.00
235	0	0	0.00	0.00
240	0	0	0.00	0.00
245	0	0	0.00	0.00
250	0	0	0.00	0.00
255	0	0	0.00	0.00

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
260	0	0	0.00	0.00
265	0	0	0.00	0.00
270	0	0	0.00	0.00
275	0	0	0.00	0.00
280	0	0	0.00	0.00
285	0	0	0.00	0.00
290	0	0	0.00	0.00
295	0	0	0.00	0.00
300	0	0	0.00	0.00
305	0	0	0.00	0.00
310	0	0	0.00	0.00
315	0	0	0.00	0.00
320	0	0	0.00	0.00
325	0	0	0.00	0.00
330	0	0	0.00	0.00
335	0	0	0.00	0.00
340	0	0	0.00	0.00
345	0	0	0.00	0.00
350	0	0	0.00	0.00
355	0	0	0.00	0.00
360	0	0	0.00	0.00
365	0	0	0.00	0.00
370	0	0	0.00	0.00
375	0	0	0.00	0.00
380	0	0	0.00	0.00
385	0	0	0.00	0.00
390	0	0	0.00	0.00
395	0	0	0.00	0.00
400	0	0	0.00	0.00
405	0	0	0.00	0.00
410	0	0	0.00	0.00
415	0	0	0.00	0.00
420	0	0	0.00	0.00
425	0	0	0.00	0.00
430	0	0	0.00	0.00
435	0	0	0.00	0.00
440	0	0	0.00	0.00
445	1	1	2.44	2.44
450	0	1	0.00	2.44
455	0	1	0.00	2.44
460	0	1	0.00	2.44
465	0	1	0.00	2.44
470	0	1	0.00	2.44
475	0	1	0.00	2.44
480	0	1	0.00	2.44
485	0	1	0.00	2.44
490	0	1	0.00	2.44
495	0	1	0.00	2.44
500	0	1	0.00	2.44
505	0	1	0.00	2.44
510	0	1	0.00	2.44
520	0	1	0.00	2.44
525	0	1	0.00	2.44
530	0	1	0.00	2.44
535	0	1	0.00	2.44

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
540	0	1	0.00	2.44
545	0	1	0.00	2.44
550	0	1	0.00	2.44
555	0	1	0.00	2.44
560	0	1	0.00	2.44
565	0	1	0.00	2.44
570	0	1	0.00	2.44
575	0	1	0.00	2.44
580	0	1	0.00	2.44
585	0	1	0.00	2.44
590	0	1	0.00	2.44
595	0	1	0.00	2.44
600	0	1	0.00	2.44
605	0	1	0.00	2.44
610	0	1	0.00	2.44
615	0	1	0.00	2.44
620	0	1	0.00	2.44
625	0	1	0.00	2.44
630	0	1	0.00	2.44
635	0	1	0.00	2.44
640	0	1	0.00	2.44
645	0	1	0.00	2.44
650	0	1	0.00	2.44
655	0	1	0.00	2.44
660	0	1	0.00	2.44
665	0	1	0.00	2.44
670	0	1	0.00	2.44
675	0	1	0.00	2.44
680	0	1	0.00	2.44
685	0	1	0.00	2.44
690	0	1	0.00	2.44
695	1	2	2.44	4.88
700	1	3	2.44	7.32
705	0	3	0.00	7.32
710	0	3	0.00	7.32
715	0	3	0.00	7.32
720	1	4	2.44	9.76
725	0	4	0.00	9.76
730	1	5	2.44	12.20
735	0	5	0.00	12.20
740	1	6	2.44	14.63
745	0	6	0.00	14.63
750	1	7	2.44	17.07
755	0	7	0.00	17.07
760	1	8	2.44	19.51
765	0	8	0.00	19.51
770	1	9	2.44	21.95
780	0	9	0.00	21.95
785	0	9	0.00	21.95
790	0	9	0.00	21.95
795	0	9	0.00	21.95
800	0	9	0.00	21.95
805	0	9	0.00	21.95
810	0	9	0.00	21.95
815	0	9	0.00	21.95

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
820	0	9	0.00	21.95
825	0	9	0.00	21.95
830	1	10	2.44	24.39
835	1	11	2.44	26.83
840	2	13	4.88	31.71
845	1	14	2.44	34.15
850	2	16	4.88	39.02
855	0	16	0.00	39.02
860	0	16	0.00	39.02
865	0	16	0.00	39.02
870	0	16	0.00	39.02
875	0	16	0.00	39.02
880	0	16	0.00	39.02
885	2	18	4.88	43.90
890	1	19	2.44	46.34
895	1	20	2.44	48.78
900	1	21	2.44	51.22
905	1	22	2.44	53.66
910	2	24	4.88	58.54
915	0	24	0.00	58.54
920	3	27	7.32	65.85
925	1	28	2.44	68.29
930	2	30	4.88	73.17
935	0	30	0.00	73.17
940	0	30	0.00	73.17
945	0	30	0.00	73.17
950	0	30	0.00	73.17
955	0	30	0.00	73.17
960	0	30	0.00	73.17
965	1	31	2.44	75.61
970	0	31	0.00	75.61
975	1	32	2.44	78.05
980	0	32	0.00	78.05
985	0	32	0.00	78.05
990	0	32	0.00	78.05
995	0	32	0.00	78.05
1000	0	32	0.00	78.05
1005	1	33	2.44	80.49
1010	1	34	2.44	82.93
1015	0	34	0.00	82.93
1020	0	34	0.00	82.93
1025	0	34	0.00	82.93
1030	1	35	2.44	85.37
1035	0	35	0.00	85.37
1040	1	36	2.44	87.80
1045	0	36	0.00	87.80
1050	0	36	0.00	87.80
1055	0	36	0.00	87.80
1060	3	39	7.32	95.12
1065	0	39	0.00	95.12
1070	0	39	0.00	95.12
1075	0	39	0.00	95.12
1080	0	39	0.00	95.12
1085	0	39	0.00	95.12
1090	0	39	0.00	95.12

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
1095	0	39	0.00	95.12
1100	1	40	2.44	97.56
1105	0	40	0.00	97.56
1110	0	40	0.00	97.56
1115	0	40	0.00	97.56
1120	0	40	0.00	97.56
1125	0	40	0.00	97.56
1130	0	40	0.00	97.56
1135	1	41	2.44	100.00

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
		Basketball, Division I		
0	0	0	0.00	0.00
5	0	0	0.00	0.00
10	0	0	0.00	0.00
15	0	0	0.00	0.00
20	0	0	0.00	0.00
25	0	0	0.00	0.00
30	0	0	0.00	0.00
35	0	0	0.00	0.00
40	0	0	0.00	0.00
45	0	0	0.00	0.00
50	0	0	0.00	0.00
55	0	0	0.00	0.00
60	0	0	0.00	0.00
65	0	0	0.00	0.00
70	0	0	0.00	0.00
75	0	0	0.00	0.00
80	0	0	0.00	0.00
85	0	0	0.00	0.00
90	0	0	0.00	0.00
95	0	0	0.00	0.00
100	0	0	0.00	0.00
105	0	0	0.00	0.00
110	0	0	0.00	0.00
115	0	0	0.00	0.00
120	0	0	0.00	0.00
125	0	0	0.00	0.00
130	0	0	0.00	0.00
135	0	0	0.00	0.00
140	0	0	0.00	0.00
145	0	0	0.00	0.00
150	0	0	0.00	0.00
155	0	0	0.00	0.00
160	0	0	0.00	0.00
165	0	0	0.00	0.00
170	0	0	0.00	0.00
175	0	0	0.00	0.00
180	1	1	2.44	2.44
185	0	1	0.00	2.44
190	0	1	0.00	2.44
195	0	1	0.00	2.44
200	0	1	0.00	2.44
205	0	1	0.00	2.44
210	0	1	0.00	2.44
215	0	1	0.00	2.44
220	0	1	0.00	2.44
225	0	1	0.00	2.44
230	0	1	0.00	2.44
235	0	1	0.00	2.44
240	0	1	0.00	2.44
245	0	1	0.00	2.44
250	0	1	0.00	2.44
255	0	1	0.00	2.44

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
260	0	1	0.00	2.44
265	0	1	0.00	2.44
270	0	1	0.00	2.44
275	0	1	0.00	2.44
280	0	1	0.00	2.44
285	0	1	0.00	2.44
290	0	1	0.00	2.44
295	3	4	7.32	9.76
300	1	5	2.44	12.20
305	1	6	2.44	14.63
310	0	6	0.00	14.63
315	0	6	0.00	14.63
320	3	9	7.32	21.95
325	0	9	0.00	21.95
330	0	9	0.00	21.95
335	0	9	0.00	21.95
340	1	10	2.44	24.39
345	4	14	9.76	34.15
350	0	14	0.00	34.15
355	5	19	12.20	46.34
360	1	20	2.44	48.78
365	8	28	19.51	68.29
370	2	30	4.88	73.17
375	0	30	0.00	73.17
380	0	30	0.00	73.17
385	1	31	2.44	75.61
390	1	32	2.44	78.05
395	0	32	0.00	78.05
400	0	32	0.00	78.05
405	0	32	0.00	78.05
410	2	34	4.88	82.93
415	0	34	0.00	82.93
420	1	35	2.44	85.37
425	2	37	4.88	90.24
430	2	39	4.88	95.12
435	0	39	0.00	95.12
440	0	39	0.00	95.12
445	0	39	0.00	95.12
450	1	40	2.44	97.56
455	0	40	0.00	97.56
460	0	40	0.00	97.56
465	1	41	2.44	100.00

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
		Basketball, Division II		
0	0	0	0.00	0.00
5	0	0	0.00	0.00
10	0	0	0.00	0.00
15	0	0	0.00	0.00
20	0	0	0.00	0.00
25	0	0	0.00	0.00
30	0	0	0.00	0.00
35	0	0	0.00	0.00
40	0	0	0.00	0.00
45	0	0	0.00	0.00
50	0	0	0.00	0.00
55	0	0	0.00	0.00
60	0	0	0.00	0.00
65	0	0	0.00	0.00
70	0	0	0.00	0.00
75	0	0	0.00	0.00
80	0	0	0.00	0.00
85	0	0	0.00	0.00
90	0	0	0.00	0.00
95	0	0	0.00	0.00
100	0	0	0.00	0.00
105	0	0	0.00	0.00
110	1	1	2.44	2.44
115	0	1	0.00	2.44
120	0	1	0.00	2.44
125	0	1	0.00	2.44
130	0	1	0.00	2.44
135	0	1	0.00	2.44
140	0	1	0.00	2.44
145	0	1	0.00	2.44
150	0	1	0.00	2.44
155	0	1	0.00	2.44
160	0	1	0.00	2.44
165	0	1	0.00	2.44
170	0	1	0.00	2.44
175	0	1	0.00	2.44
180	1	2	2.44	4.88
185	3	5	7.32	12.20
190	0	5	0.00	12.20
195	2	7	4.88	17.07
200	1	8	2.44	19.51
205	1	9	2.44	21.95
210	2	11	4.88	26.83
215	2	13	4.88	31.71
220	0	13	0.00	31.71
225	2	15	4.88	36.59
230	2	17	4.88	41.46
235	0	17	0.00	41.46
240	5	22	12.20	53.66
245	4	26	9.76	63.41
250	3	29	7.32	70.73
255	1	30	2.44	73.17

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
260	5	35	12.20	85.37
265	2	37	4.88	90.24
270	1	38	2.44	92.68
275	0	38	0.00	92.68
280	2	40	4.88	97.56
285	0	40	0.00	97.56
290	0	40	0.00	97.56
295	1	41	2.44	100.00

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
		Basketball, Division III		
0	0	0	0.00	0.00
5	0	0	0.00	0.00
10	0	0	0.00	0.00
15	0	0	0.00	0.00
20	0	0	0.00	0.00
25	0	0	0.00	0.00
30	0	0	0.00	0.00
35	0	0	0.00	0.00
40	0	0	0.00	0.00
45	0	0	0.00	0.00
50	0	0	0.00	0.00
55	0	0	0.00	0.00
60	0	0	0.00	0.00
65	0	0	0.00	0.00
70	0	0	0.00	0.00
75	0	0	0.00	0.00
80	0	0	0.00	0.00
85	0	0	0.00	0.00
90	0	0	0.00	0.00
95	0	0	0.00	0.00
100	0	0	0.00	0.00
105	0	0	0.00	0.00
110	0	0	0.00	0.00
115	0	0	0.00	0.00
120	0	0	0.00	0.00
125	0	0	0.00	0.00
130	0	0	0.00	0.00
135	0	0	0.00	0.00
140	0	0	0.00	0.00
145	0	0	0.00	0.00
150	1	1	2.44	2.44
155	0	1	0.00	2.44
160	0	1	0.00	2.44
165	0	1	0.00	2.44
170	0	1	0.00	2.44
175	0	1	0.00	2.44
180	0	1	0.00	2.44
185	0	1	0.00	2.44
190	0	1	0.00	2.44
195	0	1	0.00	2.44
200	0	1	0.00	2.44
205	0	1	0.00	2.44
210	2	3	4.88	7.32
215	0	3	0.00	7.32
220	0	3	0.00	7.32
225	0	3	0.00	7.32
230	1	4	2.44	9.76
235	1	5	2.44	12.20
240	2	7	4.88	17.07
245	2	9	4.88	21.95
250	0	9	0.00	21.95
255	0	9	0.00	21.95

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
260	0	9	0.00	21.95
265	1	10	2.44	24.39
270	1	11	2.44	26.83
275	4	15	9.76	36.59
280	2	17	4.88	41.46
285	1	18	2.44	43.90
290	3	21	7.32	51.22
295	0	21	0.00	51.22
300	2	23	4.88	56.10
305	4	27	9.76	65.85
310	1	28	2.44	68.29
315	1	29	2.44	70.73
320	3	32	7.32	78.05
325	0	32	0.00	78.05
330	1	33	2.44	80.49
335	1	34	2.44	82.93
340	0	34	0.00	82.93
345	2	36	4.88	87.80
350	1	37	2.44	90.24
355	0	37	0.00	90.24
360	2	39	4.88	95.12
365	1	40	2.44	97.56
370	0	40	0.00	97.56
375	1	41	2.44	100.00

Appendix H. Frequency - Percentage Table for Field Hockey

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
		Field Hockey, All Divisions		
0	0	0	0.00	0.00
5	0	0	0.00	0.00
10	0	0	0.00	0.00
15	0	0	0.00	0.00
20	0	0	0.00	0.00
25	0	0	0.00	0.00
30	0	0	0.00	0.00
35	0	0	0.00	0.00
40	0	0	0.00	0.00
45	0	0	0.00	0.00
50	0	0	0.00	0.00
55	0	0	0.00	0.00
60	0	0	0.00	0.00
65	0	0	0.00	0.00
70	0	0	0.00	0.00
75	0	0	0.00	0.00
80	0	0	0.00	0.00
85	0	0	0.00	0.00
90	0	0	0.00	0.00
95	0	0	0.00	0.00
100	0	0	0.00	0.00
105	0	0	0.00	0.00
110	0	0	0.00	0.00
115	0	0	0.00	0.00
120	0	0	0.00	0.00
125	0	0	0.00	0.00
130	0	0	0.00	0.00
135	0	0	0.00	0.00
140	0	0	0.00	0.00
145	0	0	0.00	0.00
150	0	0	0.00	0.00
155	0	0	0.00	0.00
160	0	0	0.00	0.00
165	0	0	0.00	0.00
170	0	0	0.00	0.00
175	0	0	0.00	0.00
180	1	1	3.13	3.13
185	0	1	0.00	3.13
190	0	1	0.00	3.13
195	0	1	0.00	3.13
200	0	1	0.00	3.13
205	0	1	0.00	3.13
210	1	2	3.13	6.25
215	0	2	0.00	6.25
220	0	2	0.00	6.25
225	0	2	0.00	6.25
230	0	2	0.00	6.25
235	0	2	0.00	6.25
240	0	2	0.00	6.25
245	0	2	0.00	6.25
250	1	3	3.13	9.38
255	1	4	3.13	12.50

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
260	1	5	3.13	15.63
265	0	5	0.00	15.63
270	0	5	0.00	15.63
275	0	5	0.00	15.63
280	0	5	0.00	15.63
285	1	6	3.13	18.75
290	1	7	3.13	21.88
295	1	8	3.13	25.00
300	1	9	3.13	28.13
305	1	9	0.00	28.13
310	1	9	0.00	28.13
315	1	11	3.13	31.25
321	2	12	6.25	37.50
325	0	12	0.00	37.50
330	0	12	0.00	37.50
335	0	12	0.00	37.50
340	1	13	3.13	40.63
345	0	13	0.00	40.63
350	0	13	0.00	40.63
355	2	15	6.25	46.88
360	0	15	0.00	46.88
365	1	16	3.13	50.00
370	0	16	0.00	50.00
375	2	18	6.25	56.25
381	2	20	6.25	62.50
385	3	23	9.38	71.88
391	3	26	9.38	81.25
395	0	26	0.00	81.25
400	0	26	0.00	81.25
405	0	26	0.00	81.25
410	1	27	3.13	84.38
415	1	28	3.13	87.50
421	2	30	6.25	93.75
425	0	30	0.00	93.75
431	2	32	6.25	100.00

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
		Field Hockey, Division I		
0	0	0	0.00	0.00
5	0	0	0.00	0.00
10	0	0	0.00	0.00
15	0	0	0.00	0.00
20	0	0	0.00	0.00
25	0	0	0.00	0.00
30	0	0	0.00	0.00
35	0	0	0.00	0.00
40	0	0	0.00	0.00
45	0	0	0.00	0.00
50	0	0	0.00	0.00
55	0	0	0.00	0.00
60	0	0	0.00	0.00
65	0	0	0.00	0.00
70	1	1	3.13	3.13
75	0	1	0.00	3.13
80	0	1	0.00	3.13
85	1	2	3.13	6.25
91	2	4	6.25	12.50
95	1	4	0.00	12.50
100	1	5	3.13	15.63
105	0	5	0.00	15.63
110	0	5	0.00	15.63
115	2	7	6.25	21.88
121	2	9	6.25	28.13
125	2	11	6.25	34.38
130	1	12	3.13	37.50
135	1	13	3.13	40.63
140	0	13	0.00	40.63
145	2	15	6.25	46.88
151	3	18	9.38	56.25
155	2	20	6.25	62.50
161	6	26	18.75	81.25
165	3	29	9.38	90.63
170	0	29	0.00	90.63
175	1	31	3.13	93.75
180	0	30	0.00	93.75
185	2	32	6.25	100.00

MIDPOINT	FREQ	CUM. FREQ	%	CUM %
		Field Hockey, Division III		
0	0	0	0.00	0.00
5	0	0	0.00	0.00
10	0	0	0.00	0.00
15	0	0	0.00	0.00
20	0	0	0.00	0.00
25	0	0	0.00	0.00
30	0	0	0.00	0.00
35	0	0	0.00	0.00
40	0	0	0.00	0.00
45	0	0	0.00	0.00
50	0	0	0.00	0.00
55	0	0	0.00	0.00
60	0	0	0.00	0.00
65	0	0	0.00	0.00
70	0	0	0.00	0.00
75	0	0	0.00	0.00
80	0	0	0.00	0.00
85	0	0	0.00	0.00
90	0	0	0.00	0.00
95	0	0	0.00	0.00
100	0	0	0.00	0.00
105	0	0	0.00	0.00
110	1	1	3.13	3.13
115	0	1	0.00	3.13
120	0	1	0.00	3.13
125	1	2	3.13	6.25
130	0	2	0.00	6.25
135	0	2	0.00	6.25
140	0	2	0.00	6.25
145	0	2	0.00	6.25
150	0	2	0.00	6.25
155	0	2	0.00	6.25
161	3	5	9.38	15.63
165	0	5	0.00	15.63
171	3	8	9.38	25.00
175	1	9	3.13	28.13
180	1	9	0.00	28.13
185	1	9	0.00	28.13
190	1	9	0.00	28.13
195	2	11	6.25	34.38
201	2	13	6.25	40.63
205	1	14	3.13	43.75
210	1	15	3.13	46.88
215	2	17	6.25	53.13
220	1	18	3.13	56.25
225	4	22	12.50	68.75
231	3	25	9.38	78.13
235	1	26	3.13	81.25
240	0	26	0.00	81.25
245	4	30	12.50	93.75
250	0	30	0.00	93.75
255	2	32	6.25	100.00

Appendix I. Question Ranks - Basketball

Basketball, All Divisions

% - PTS	OBS	RANK	DIV	QUES	%
	1	1131	BT	Q6	
	2	1097	BT	Q11	
87.80	3	1056	BT	Q9	90
1052	4	1055	BT	Q10	
	5	1054	BT	Q1	
85.37	6	1035	BT	Q5	85
1027	7	1024	BT	Q4	
80.49	8	1007	BT	Q12	80
1002	9	1001	BT	Q7	
	10	972	BT	Q40	
	11	962	BT	Q41	
	12	930	BT	Q20	
	13	925	BT	Q2	
	14	922	BT	Q24	
	15	920	BT	Q21	
	16	919	BT	Q13	
	17	917	BT	Q28	
	18	908	BT	Q26	
	19	906	BT	Q25	
	20	901	BT	Q30	
	21	897	BT	Q22	
	22	893	BT	Q27	

OBS	RANK	DIV	QUES
23	888	BT	Q29
24	885	BT	Q23
25	883	BT	Q31
26	847	BT	Q18
27	846	BT	Q38
28	844	BT	Q16
29	840	BT	Q39
30	838	BT	Q19
31	834	BT	Q36
32	828	BT	Q37
33	767	BT	Q32
34	760	BT	Q33
35	746	BT	Q3
36	735	BT	Q14
37	726	BT	Q8
38	716	BT	Q17
39	699	BT	Q34
40	691	BT	Q35
41	441	BT	Q15

Basketball, Division I

% - PTS	OBS	RANK	DIV	QUES	%
	1	463	B1	Q6	
	2	452	B1	Q11	
90.24	3	429	B1	Q9	90
422	4	428	B1	Q10	
	5	425	B1	Q1	
	6	425	B1	Q5	
85.37	7	420	B1	Q4	85
417					
78.05	8	410	B1	Q7	80
397	9	410	B1	Q12	
	10	390	B1	Q40	
	11	387	B1	Q41	
	12	370	B1	Q24	
	13	368	B1	Q13	
	14	367	B1	Q30	
	15	366	B1	Q20	
	16	366	B1	Q21	
	17	366	B1	Q25	
	18	365	B1	Q31	
	19	364	B1	Q2	
	20	363	B1	Q22	
	21	363	B1	Q23	
	22	360	B1	Q28	
	23	353	B1	Q26	

OBS	RANK	DIV	QUES
24	353	B1	Q27
25	353	B1	Q29
26	353	B1	Q38
27	353	B1	Q39
28	346	B1	Q18
29	346	B1	Q19
30	346	B1	Q36
31	346	B1	Q37
32	342	B1	Q16
33	321	B1	Q14
34	320	B1	Q32
35	318	B1	Q33
36	307	B1	Q3
37	300	B1	Q17
38	295	B1	Q35
39	293	B1	Q8
40	293	B1	Q34
41	180	B1	Q15

Basketball, Division II

% - PTS	OBS	RANK	DIV	QUES	%
	1	293	B2	Q6	
	2	280	B2	Q10	
	3	278	B2	Q11	
90.24	4	268	B2	Q1	90
262	5	265	B2	Q9	
	6	264	B2	Q40	
	7	262	B2	Q7	
	8	262	B2	Q12	
	9	261	B2	Q4	
85.37	10	261	B2	Q5	85
257	11	258	B2	Q41	80
	12	253	B2	Q20	
	13	252	B2	Q28	
	14	251	B2	Q26	
	15	250	B2	Q21	
	16	246	B2	Q24	
	17	246	B2	Q30	
	18	245	B2	Q22	
	19	243	B2	Q29	
	20	242	B2	Q2	
	21	242	B2	Q27	
	22	239	B2	Q23	
	23	238	B2	Q25	
	24	238	B2	Q31	

OBS	RANK	DIV	QUES
25	232	B2	Q18
26	231	B2	Q13
27	225	B2	Q19
28	224	B2	Q16
29	217	B2	Q38
30	215	B2	Q36
31	211	B2	Q39
32	209	B2	Q37
33	206	B2	Q32
34	200	B2	Q33
35	196	B2	Q3
36	194	B2	Q34
37	186	B2	Q8
38	186	B2	Q17
39	186	B2	Q35
40	179	B2	Q14
41	110	B2	Q15

Basketball, Division III

% - PTS	OBS	RANK	DIV	QUES	%
	1	375	B3	Q6	
	2	367	B3	Q11	
90.24	3	362	B3	Q8	90
347	4	361	B3	Q1	
	5	349	B3	Q5	
	6	347	B3	Q10	
82.93	7	343	B3	Q4	
332	8	335	B3	Q12	
80.49	9	329	B3	Q7	80
332					
	10	320	B3	Q13	
	11	319	B3	Q2	
	12	318	B3	Q40	
	13	317	B3	Q41	
	14	311	B3	Q20	
	15	306	B3	Q24	
	16	305	B3	Q28	
	17	304	B3	Q21	
	18	304	B3	Q26	
	19	302	B3	Q25	
	20	298	B3	Q27	
	21	292	B3	Q29	
	22	289	B3	Q22	
	23	288	B3	Q30	

OBS	RANK	DIV	QUES
24	283	B3	Q23
25	280	B3	Q31
26	278	B3	Q16
27	276	B3	Q38
28	276	B3	Q39
29	273	B3	Q36
30	273	B3	Q37
31	269	B3	Q18
32	267	B3	Q19
33	247	B3	Q9
34	243	B3	Q3
35	242	B3	Q33
36	241	B3	Q32
37	235	B3	Q14
38	230	B3	Q17
39	212	B3	Q34
40	210	B3	Q35
41	151	B3	Q15

Appendix J. Question Ranks - Field Hockey

Field Hockey, All Divisions

% - PTS	OBS	RANK	DIV	QUES	%
	1	432	HT	Q6	
	2	428	HT	Q11	
87.50	3	421	HT	Q1	90
412	4	418	HT	Q21	
	5	415	HT	Q22	
84.38	6	411	HT	Q10	85
407					
	7	391	HT	Q27	
81.25	8	389	HT	Q2	80
387	9	388	HT	Q9	
	10	385	HT	Q12	
	11	384	HT	Q17	
	12	384	HT	Q28	
	13	379	HT	Q7	
	14	378	HT	Q23	
	15	377	HT	Q5	
	16	373	HT	Q24	
	17	366	HT	Q31	
	18	357	HT	Q4	
	19	356	HT	Q32	
	20	340	HT	Q13	
	21	321	HT	Q25	

OBS	RANK	DIV	QUES
22	318	HT	Q19
23	314	HT	Q26
24	302	HT	Q3
25	295	HT	Q8
26	288	HT	Q16
27	283	HT	Q14
28	262	HT	Q20
29	254	HT	Q29
30	252	HT	Q30
31	208	HT	Q18
32	179	HT	Q15

Field Hockey, Division I

% - PTS	OBS	RANK	DIV	QUES	%
	1	186	H1	Q6	
	2	183	H1	Q11	
	3	177	H1	Q1	
	4	165	H1	Q10	
90.63	5	164	H1	Q21	90
162	6	163	H1	Q12	
	7	162	H1	Q22	
	8	161	H1	Q9	
	9	161	H1	Q27	
81.25	10	160	H1	Q5	85
157	11	159	H1	Q2	80
	12	159	H1	Q28	
	13	155	H1	Q4	
	14	154	H1	Q7	
	15	149	H1	Q31	
	16	148	H1	Q17	
	17	148	H1	Q23	
	18	147	H1	Q24	
	19	147	H1	Q32	
	20	137	H1	Q13	
	21	129	H1	Q3	
	22	127	H1	Q8	
	23	124	H1	Q19	
	24	120	H1	Q16	

OBS	RANK	DIV	QUES
25	119	H1	Q25
26	117	H1	Q26
27	114	H1	Q14
28	100	H1	Q20
29	92	H1	Q29
30	91	H1	Q30
31	84	H1	Q18
32	71	H1	Q15

Field Hockey, Division III

% - PTS	OBS	RANK	DIV	QUES	%
	1	280	H3	Q23	
	2	254	H3	Q21	
	3	253	H3	Q22	
93.75	4	246	H3	Q6	90
242	5	246	H3	Q10	
	6	245	H3	Q11	
81.25	7	244	H3	Q1	85
232					80
	8	236	H3	Q17	
	9	230	H3	Q2	
	10	230	H3	Q27	
	11	227	H3	Q9	
	12	226	H3	Q24	
	13	225	H3	Q7	
	14	225	H3	Q28	
	15	222	H3	Q12	
	16	217	H3	Q5	
	17	217	H3	Q31	
	18	209	H3	Q32	
	19	203	H3	Q13	
	20	202	H3	Q4	
	21	202	H3	Q25	
	22	197	H3	Q26	
	23	194	H3	Q19	

OBS	RANK	DIV	QUES
24	173	H3	Q3
25	169	H3	Q14
26	168	H3	Q8
27	168	H3	Q16
28	162	H3	Q20
29	162	H3	Q29
30	161	H3	Q30
31	124	H3	Q18
32	108	H3	Q15

Appendix K. Instructions to Control Data Evaluators

The following is a grouping of data that depicts two stages in a hypothetical basketball season. There are twenty teams being considered for a Top Twenty Poll. In the spaces provided to the left of the team name, please place a number corresponding to the position in which that team should be ranked. In this poll, "1" is the highest rank and "20" is the lowest.

Below, you will find an explanation for each of the columns in the data set.

1. *Team Name.*

2. *Record vs Common Opponents*: Assume that the only common opponents are from within this group. *NOTE*: Not all teams play all of the other teams. In other words, the record in this column may not equal the total season record. In this column, as in all those dealing with wins and losses, the first number is representative of the number of wins, the second represents losses.

3. *Scores vs Common Opponents*: The value in each cell is an expression of scoring margin among the games between common opponents. A value of **+28** means the total points scored by Team X was 28 more than the total points Team X gave up, or allowed. Similarly, a value of **-52** means the total points scored by Team Y was 52 less than the total points Team Y allowed. The scoring margin is only recorded for games among common opponents.

4. *Strength of Schedule, beginning of season*: Index based on 100 schedules; a value of 14 in the cell means that Team X had the 14th toughest schedule. This value is determined prior to the beginning of the season.

5. *Strength of Schedule, during the season*: Index based on 100 schedules. The assumption is that the "strength" changes as teams move in and out of the poll.

6. *Record vs Top 20*: Win/loss record against teams currently in the Top Twenty Poll.

7. *"Credible Losses:"* A loss to a higher ranked team by less than 5 points.

8. *Week's Record*: The win/loss record for the current week. Current week is the week under consideration for the poll.

9. *Season's Record*: The win/loss record for the season up to the point of consideration.

10. *Record vs 1-10*: The win/loss record for teams ranked in the top 20 in positions 1 through 10.

11. *Record vs 11-20*: The win/loss record for teams ranked in the top 20 in positions 11 through 20.

12. *Offensive Statistic*: The value in each cell is an expression representing the difference in field goal percentage made and field goal percentage allowed, expressed in thousandths of a point. For example, Team X has a field goal percentage of **.481**, and has allowed a field goal percentage of **.441**. The difference, **.040**, is the value in the cell. A negative value means Team Y gave up more percentage points than it made.

Please put your name in the space provided so I can credit you in my thesis.
Thank you for your assistance.

Raz

Appendix L. Data Used to Build the Control Data Sets

NBA Data, 1977 Season[1]

P	TEAM	W	L	%	FG %+	%P	FG %-	FG %δ	FG %δP
4	Philadelphia	33	20	.623	.479	6(T)	.441	+.038	1
13	Boston	26	27	.491	.446	18	.452	-.006	14
16	NY Knicks	24	29	.453	.479	6(T)	.469	+.010	7
20	Buffalo	19	33	.365	.449	17	.475	-.026	20
21	NY Nets	17	36	.321	.430	22	.461	-.031	22
6	Washington	30	22	.577	.466	10	.459	+.007	8
7	Houston	29	23	.558	.487	2	.470	+.017	6
9	Cleveland	28	24	.538	.443	19(T)	.443	+.000	12
10	San Antonio	29	25	.537	.476	9	.486	-.010	15(T)
17	New Orleans	24	30	.444	.451	16	.450	+.001	11
19	Atlanta	22	34	.393	.458	14	.479	-.021	18
2	Denver	34	19	.642	.481	4(T)	.463	+.018	4(T)
5	Detroit	32	23	.582	.490	1	.472	+.018	4(T)
12	Kansas City	28	28	.500	.459	12(T)	.469	-.010	15(T)
15	Indiana	25	30	.455	.443	19(T)	.465	-.022	19
18	Chicago	23	32	.418	.443	19(T)	.471	-.028	21
22	Milwaukee	17	42	.288	.460	11	.479	-.019	17
1	Los Angeles	35	19	.648	.482	3	.446	+.036	2
3	Portland	35	21	.625	.481	4(T)	.453	+.028	3
8	Golden State	30	24	.556	.477	8	.473	+.004	10
11	Seattle	29	26	.527	.453	15	.458	-.005	13
14	Phoenix	25	27	.481	.459	12(T)	.453	+.006	9

[1]From Friedman (1978).

Scoring Margin, NCAA 1986 Season[1]

POS	TEAM	SCO
1	Cleveland State	20.9
2	North Carolina	18.8
3	Navy	15.3
4	Syracuse	15.1
5	Memphis State	15.0
6	Notre Dame	14.8
7	Georgetown	14.6
8	Michigan	13.6
9	Kentucky	13.5
10	Oklahoma	13.2
11	Duke	13.0
12	Kansas	13.0
13	Michigan State	12.9
14	Georgia Tech	12.2
15	Texas Christian	11.9
16[2]	St. John's(NY)	11.8

[1]From NCAA News, March 12, 1986

[2]Positions 17-20 were made up for the control data.

Appendix M. Preliminary Models and Resulting Polls

On the following pages are the fifteen variations of the initial model and the polls resulting from each equation and the second control data set. It should be noted, that the following variables were integers instead of indices: CL, and SCO.

PTS = R20 + SCO-SOSB-SOSD + RCO-CL + RWK + RSN + R10 + R11 + STAT

POS	TEAM	PTS
1	Randolph	21.03
2	Davidson	4.71
3	Derring	0.54
4	Holden	-0.31
5	Williams	-4.13
6	McBryde	-6.92
7	Whittemore	-10.81
8	Hutcheson	-16.78
9	Pamplin	-32.36
10	Sandy	-38.94
11	Price	-40.66
12	Smyth	-43.53
13	Agnew	-44.42
14	Norris	-56.15
15	Patton	-65.46
16	Wallace	-68.00
17	Henderson	-69.66
18	Cowgill	-92.11
19	Robeson	-93.18
20	Seitz	-134.90

PTS = R20 + SCO-SOSD + RCO-CL + RWK + RSN + R10 + R11 + STAT

POS	TEAM	PTS
1	Randolph	23.027
2	Derring	10.539
3	Davidson	5.708
4	Holden	4.686
5	Whittemore	4.190
6	Williams	3.873
7	McBryde	2.083
8	Hutcheson	-1.778
9	Sandy	-7.940
10	Pamplin	-10.361
11	Smyth	-13.533
12	Price	-13.661
13	Agnew	-15.424
14	Norris	-21.152
15	Patton	-25.464
16	Henderson	-28.665
17	Wallace	-30.005
18	Cowgill	-36.114
19	Robeson	-41.182
20	Seitz	-59.901

PTS = R20 + SCO-SOSB + RCO-CL + RWK + RSN + R10 + R11 + STAT

POS	TEAM	PTS
1	Randolph	22.027
2	Derring	9.539
3	Davidson	7.708
4	Holden	4.686
5	Whittemore	4.190
6	Williams	3.873
7	Hutcheson	2.222
8	McBryde	-0.917
9	Sandy	-7.940
10	Pamplin	-12.361
11	Smyth	-13.533
12	Price	-13.661
13	Agnew	-15.424
14	Norris	-18.152
15	Patton	-24.464
16	Wallace	-29.005
17	Henderson	-29.665
18	Cowgill	-38.114
19	Robeson	-43.182
20	Seitz	-59.901

PTS = SCO-SOSD + RCO-CL + RWK + RSN + R10 + R11 + STAT

POS	TEAM	PTS
1	Randolph	22.472
2	Derring	9.847
3	Davidson	5.447
4	Holden	4.486
5	Williams	3.540
6	Whittemore	3.366
7	McBryde	1.662
8	Hutcheson	-2.445
9	Sandy	-8.690
10	Pamplin	-10.861
11	Smyth	-14.133
12	Price	-14.528
13	Agnew	-15.674
14	Norris	-22.009
15	Patton	-26.114
16	Henderson	-29.040
17	Wallace	-30.505
18	Cowgill	-36.960
19	Robeson	-41.515
20	Seitz	-60.714

PTS = R20 + SCO-SOSD + RCO-CL + RWK + RSN + STAT

POS	TEAM	PTS
1	Randolph	22.194
2	Derring	9.145
3	Davidson	5.240
4	Holden	4.353
5	Williams	3.262
6	Whittemore	2.490
7	McBryde	1.254
8	Hutcheson	-3.143
9	Sandy	-9.540
10	Pamplin	-10.961
11	Smyth	-14.200
12	Price	-15.461
13	Agnew	-15.674
14	Norris	-21.152
15	Patton	-26.662
16	Henderson	-29.427
17	Wallace	-31.000
18	Cowgill	-37.914
19	Robeson	-41.817
20	Seitz	-61.687

PTS = R10 + R11 + SCO-SOSD + RCO -CL + RWK + RSN + STAT

POS	TEAM	PTS
1	Randolph	22.472
2	Derring	9.847
3	Davidson	5.447
4	Holden	4.486
5	Williams	3.540
6	Whittemore	3.366
7	McBryde	1.662
8	Hutcheson	-2.445
9	Sandy	-8.690
10	Pamplin	-10.861
11	Smyth	-14.133
12	Price	-14.528
13	Agnew	-15.674
14	Norris	-22.009
15	Patton	-26.114
16	Henderson	-29.040
17	Wallace	-30.505
18	Cowgill	-36.960
19	Robeson	-41.515
20	Seitz	-60.714

PTS = R20 + SCO-SOSD + RSN

POS	TEAM	PTS
1	Randolph	21.103
2	Holden	11.125
3	Davidson	10.652
4	Derring	7.415
5	McBryde	5.165
6	Williams	4.586
7	Whittemore	0.982
8	Hutcheson	-4.805
9	Sandy	-10.808
10	Pamplin	-11.235
11	Smyth	-13.818
12	Agnew	-14.595
13	Price	-15.640
14	Norris	-21.766
15	Henderson	-26.125
16	Wallace	-26.145
17	Patton	-26.312
18	Cowgill	-39.417
19	Robeson	-39.455
20	Seitz	-61.506

PTS = R10 + R11 + SCO-SOSD + RSN

POS	TEAM	PTS
1	Randolph	21.381
2	Holden	11.258
3	Davidson	10.859
4	Derring	8.117
5	McBryde	5.573
6	Williams	4.864
7	Whittemore	1.858
8	Hutcheson	-4.107
9	Sandy	-9.958
10	Pamplin	-11.135
11	Smyth	-13.751
12	Agnew	-14.595
13	Price	-14.707
14	Norris	-22.623
15	Wallace	-25.650
16	Henderson	-25.738
17	Patton	-25.764
18	Cowgill	-38.463
19	Robeson	-39.153
20	Seitz	-60.533

PTS = R20-SOSD + RCO-CL + RWK + RSN + STAT

POS	TEAM	PTS
1	Randolph	1.294
2	Derring	-5.955
3	Davidson	-7.660
4	Williams	-8.538
5	McBryde	-9.046
6	Holden	-10.947
7	Whittemore	-12.110
8	Hutcheson	-16.143
9	Pamplin	-18.861
10	Price	-25.561
11	Sandy	-28.340
12	Smyth	-29.200
13	Agnew	-29.274
14	Norris	-35.952
15	Patton	-40.162
16	Henderson	-42.427
17	Wallace	-42.900
18	Cowgill	-51.114
19	Robeson	-51.617
20	Seitz	-73.887

PTS = R10 + R11-SOSD + RCO-CL + RWK + RSN + STAT

POS	TEAM	PTS
1	Randolph	1.572
2	Derring	-5.253
3	Davidson	-7.453
4	Williams	-8.260
5	McBryde	-8.638
6	Holden	-10.814
7	Whittemore	-11.234
8	Hutcheson	-15.445
9	Pamplin	-18.761
10	Price	-24.653
11	Sandy	-27.490
12	Smyth	-29.133
13	Agnew	-29.274
14	Norris	-36.809
15	Patton	-39.614
16	Henderson	-42.040
17	Wallace	-42.405
18	Cowgill	-50.160
19	Robeson	-51.315
20	Seitz	-72.914

PTS = R20-SOSD + RWK + RSN + STAT

POS	TEAM	PTS
1	Randolph	0.739
2	Davidson	-1.921
3	Holden	-4.147
4	McBryde	-4.467
5	Derring	-6.647
6	Williams	-6.871
7	Whittemore	-12.934
8	Hutcheson	-16.810
9	Pamplin	-18.361
10	Price	-24.395
11	Agnew	-27.524
12	Smyth	-28.800
13	Sandy	-29.090
14	Norris	-35.809
15	Wallace	-37.400
16	Henderson	-38.802
17	Patton	-38.812
18	Robeson	-48.950
19	Cowgill	-51.960
20	Seitz	-73.700

PTS = R10 + R11-SOSD + RWK + RSN + STAT

POS	TEAM	PTS
1	Randolph	1.017
2	Davidson	-1.714
3	Holden	-4.014
4	McBryde	-4.059
5	Derring	-5.945
6	Williams	-6.593
7	Whittemore	-12.058
8	Hutcheson	-16.112
9	Pamplin	-18.261
10	Price	-24.495
11	Agnew	-27.524
12	Sandy	-28.240
13	Smyth	-28.733
14	Norris	-36.666
15	Wallace	-36.905
16	Patton	-38.264
17	Henderson	-38.415
18	Robeson	-48.648
19	Cowgill	-51.006
20	Seitz	-72.727

PTS = R20-SOSD + RWK + RSN

POS	TEAM	PTS
1	Randolph	0.703
2	Davidson	-1.915
3	Holden	-4.175
4	McBryde	-4.468
5	Derring	-6.685
6	Williams	-6.881
7	Whittemore	-12.951
8	Hutcheson	-16.805
9	Pamplin	-18.335
10	Price	-25.407
11	Agnew	-27.528
12	Smyth	-28.818
13	Sandy	-29.108
14	Norris	-35.816
15	Wallace	-37.378
16	Henderson	-38.792
17	Patton	-38.812
18	Robeson	-48.922
19	Cowgill	-51.950
20	Seitz	-73.706

PTS = R10 + R11-SOSD + RWK + RSN

POS	TEAM	PTS
1	Randolph	0.981
2	Davidson	-1.708
3	Holden	-4.042
4	McBryde	-4.060
5	Derring	-5.983
6	Williams	-6.603
7	Whittemore	-12.075
8	Hutcheson	-16.107
9	Pamplin	-18.235
10	Price	-24.474
11	Agnew	-27.528
12	Sandy	-28.258
13	Smyth	-28.751
14	Norris	-36.673
15	Wallace	-36.883
16	Patton	-38.264
17	Henderson	-38.405
18	Robeson	-48.620
19	Cowgill	-50.996
20	Seitz	-72.733

PTS = R20 + SCO-SOSB + RCO-CL + RWK + RSN + STAT

POS	TEAM	PTS
1	Randolph	21.194
2	Derring	8.145
3	Davidson	7.240
4	Holden	4.353
5	Williams	3.262
6	Whittemore	2.490
7	Hutcheson	0.857
8	McBryde	-1.746
9	Sandy	-9.540
10	Pamplin	-12.961
11	Smyth	-14.200
12	Price	-15.461
13	Agnew	-15.674
14	Norris	-18.152
15	Patton	-25.662
16	Wallace	-30.000
17	Henderson	-30.427
18	Cowgill	-39.914
19	Robeson	-43.817
20	Seitz	-61.687

The vita has been removed from the scanned document