CHAPTER FOUR

RESULTS AND DISCUSSION

4.1) INTRODUCTION

The goal of this study was to compare the results obtained when a set of variables obtained from certain test results were imputed into the discriminant analysis method, and then when this same set of variables were imputed into logistical regression. As was noted in the previous chapter, the method of discriminant analysis has traditionally been used as a sole basis of test result interpretation in South Africa. Logistical regression is in turn more commonly in use in the old East Bloc countries.

There are other equally effective methods however, with logistical regression being an example of such. Therefore, the need arose to examine the two, and to see what results were obtained by both, hence the origin of this study. Since the overall drive in South Africa is that of talent identification in general and in with particular focus on the previously disadvantaged populations, all the while remaining competitive in the international sporting arena, it is vitally important that all avenues be investigated, so as to ascertain whether we as a country are adopting and following the most effective options, all the while bearing in mind the time and money factors involved.

These variables were obtained by testing groups of under-11 rugby players that belonged to known groups of talent, i.e.: talented and less talented rugby players. A complete description of the test battery followed can be examined in chapter three. In this chapter, the descriptive statistics of the 2 groups will be examined in detail.

The results obtained from the tests done on the two groups will be examined, with a particular focus on the results obtained via discriminant analysis and logistical regression. The format will be predominantly in table form, with a written description that will follow.

4.2) COMPARISON OF DESCRIPTIVE STATISTICS OF THE 2 GROUPS

Before representing the table containing the raw data of the descriptive statistics obtained in the test battery that was followed, it is important to discuss these values. For the purposes of this explanation, the following values were observed and discussed and notated:

Table 2) Key explaining the symbols used in the notation of the findings of this study

VALUE OR OBSERVATION	NOTATION
1. Mean	$\bar{X}$
2. Standard deviation	S
3. Minimum and Maximum Value	Min and Max
4. d-value	d-value

The d-value represents the effect size, which is the standardised means difference between two groups that are compared to one another. It is a formula derived from the difference between the two means of the two groups in question, divided by the maximum standard deviation found in either of the groups in question. It is represented in the following way:

d value = (Mean 1 – Mean 2) ÷ Standard Deviation Max

OR, in notation it can be read as:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{S_{max}}$$ (Steyn, 2000)

where:

$\bar{X}_1$ is the mean if the first group and $\bar{X}_2$ is the mean of the second group, and where S_{max} is representative of the maximum standard deviation that occurs between the two groups under evaluation.

An important note is the guidelines to interpret "effect sizes" or "practical significances." The guideline value for a small effect or significance is d = 0.2, for a medium effect or significance is d = 0.5, and for a large effect or significance is d = 0.8 (therefore practically significant, or of practical significance).

Note that d can be negative when $\bar{X}_1$ is smaller than $\bar{X}_2$, therefore the guidelines are for the absolute values of d (i.e. ignoring the sign).

What follows below is the table containing the raw data of the descriptive statistics obtained in the test battery that was followed. The X, S, min/max and d-values were observed and discussed and notated

Table 3) Descriptive statistics and significant differences of the SARFU and MSP groups with regards to rugby-specific skills, motor-physical abilities and anthropometrical data.

	MSP		n = 40		SARFU		n = 43		
VARIABLE	X	S	Min	Max	X	S	Min	Max	d-value
Body Mass (kg)	56.4	10.36	30.4	81.2	56	9.78	40.3	83.9	0.04
Triceps SF (mm)	21.6	6.75	8	32.4	9.6	4.8	4.4	22	1.8
Subscapular SF (mm)	22.72	10.92	6	43.6	8.98	3.89	5.5	23.4	1.26
Supra-spinal SF (mm)	25.4	12	4	46.6	8.74	8.53	4	51.1	1.39
Medial Calf SF (mm)	19.96	5.2	8	32.5	10.39	5.39	4.8	25.6	1.78
Upper Arm Girth (cm)	28.32	2.87	22	33.6	27.49	4.86	2.5	33	0.17
Calf Girth (cm)	34.38	2.97	27	42	33.34	2.92	28.9	41.5	0.35
Humerus Breadth (cm)	6.87	0.49	6	8.5	6.8	1.57	0.8	8.6	0.04
Femur Breadth (cm)	9.83	0.72	8.3	12	9.86	0.74	8.6	11.6	-0.04
Passing for Distance (m)	11.75	2.16	7.5	17	16.18	2.32	10.2	21	-1.91
Passing for Accuracy 7m (pts)	16.6	5.03	4	25	14.47	2.47	8	18	0.42
Passing for Accuracy 4m (pts)	2.47	1.9	0	7	3.83	2.2	0	7	-0.62
Running and Catching	14.79	4.85	0	20	18.83	1.54	14	20	-0.83
Kick for Distance (m)	23.78	5.58	10.3	33.4	30.52	5.12	19.8	43.5	-1.2
Kick-off for Distance (m)	22.89	5.73	10	32.4	32.54	7.42	17	45.5	-1.3
Speed (sec)	9.98	9.4	7	70	6.7	0.54	5.9	8.1	0.34
Flexed Armhang (sec)	22.53	14.44	1.4	68	43.3	20.36	0	83	-1.02
Vertical Jump (cm)	25.38	7.43	8	45	41.3	6.56	25	51	-2.14
Agility Run (sec)	8.26	2.82	0	11.9	7.11	0.45	6.3	8.1	0.41
Speed Endurance 1 (sec)	11.56	0.98	9.6	14.5	9.43	0.44	8.4	10.7	2.17
Speed Endurance 2 (sec)	11.67	0.98	9.6	14.4	9.59	0.55	8.8	11.2	1.1
Speed Endurance 3 (sec)	12.16	1.04	9.9	14.4	9.67	0.46	9	10.8	2.39
Speed Endurance 4 (sec)	11.73	2.79	0	15	9.68	0.45	8.9	11	0.73
Speed Endurance 5 (sec)	11.89	2.98	0	16.6	9.71	0.52	9	11.5	0.73
Speed Endurance 6 (sec)	11.43	2.8	0	16.3	9.65	0.47	8.9	11	0.64

d>0.8=Great Effect
d>0.5=Medium Effect
d>0.2=Small Effect (Hare, 1999)

4.2.1) Characteristics of Table 3

The characteristics and comparisons between the two groups will be discussed according to the following headings:

- Rugby-specific skills
- Motor-physical abilities
- Anthropometrical data

4.2.1.1) Sports-specific skills

In the above category, the SARFU group performs better in five out of the six tests (Table 3). The only test where the MSP group outdoes the SARFU group is in the forward pass for accuracy over 7m. When the significance or effect size is evaluated, it can be seen that that in all but 1 test, the SARFU group obtains a significance of 0.8 or higher, indicating a large significance.

The test that ranks as medium effect for the SARFU group is the forward pass for accuracy, in which they obtain an effect size or significance of –0.62. It is interesting to note that in the test in which the MSP performs better than the SARFU group, the effect size ranks as a small to medium effect, obtaining a d-value of 0.42. When the results of the above category are examined, certain realities boldly stand out.

As is known from previous discussions, the SARFU group is the "more talented" group. When this is considered, it can be seen that sports-specific skills rank highly in terms of distinguishing between the more talented and the less talented players, with extremely high significances in terms of the different skills that are needed to be successful in the game of rugby.

As a side note, this emphasizes the fact that skills need to be properly and exhaustively taught at an early age, for the proper effect to be seen at the same age, but more importantly later. For South Africa to remain competitive on the world stage, these skills need to be taught across the board, to all population groups, starting at a young age when the most important learning phases in terms of skills can and should take place.

4.2.1.2) Physical-motor abilities

As was mentioned in the discussion of the previous section (Table 3), it can be seen from the results obtained in this section that the SARFU group are indeed the more talented group, and it shows in the results obtained in the physical-motor abilities. SARFU outperformed the MSP group in all of the above tests.

When the d-values and therefore the significances are evaluated, the following is found:

There is a very high significance in five of the categories, namely flexed arm-hang, vertical jump, speed endurance 1, 2 and 3. The significance or d-value obtained is 1.02 or higher. There is a medium to high significance in three of the categories, namely speed endurance 4, 5 and 6, where the lowest d-value is 0.64. There are two categories in which the d-value is low, namely speed and agility run, where the d-values are 0.34 and 0.41 respectively, indicating a small to medium effect or significance.

What is significant from the results in this category is that SARFU outperforms MSP in all of the tests. This once again illustrates that the athletes have been sorted into the correct groups, but also that these skills or abilities are needed to be successful in the game of rugby, albeit that some skills are more significant than others.

The skills that appear, according to the test findings, to have more significance in terms of success are flexed arm-hang, vertical jump and the speed endurance tests. This does not necessarily mean that the others are not needed, but could indicate that they may not be as important as the above-mentioned skills.

This could emphasize that even at this young age certain motor-physical abilities need to be present, or at least visible or taught (trained). Therefore, a certain level of sports-specific training in terms of the above abilities would need to be performed. What should be noted, however, is that these are not necessarily abilities inherent in rugby only, but rather general skills trained to all participants in all sports.

4.2.1.3) Anthropometrical data

The results obtained in the above tests are more evenly mixed (Table 3), with SARFU outperforming MSP slightly. It is important however, that the results are carefully interpreted so as to get a true reflection of the situation that exists. Firstly, the skin fold results have been interpreted on that basis of the highest score not necessarily meaning the better result. Therefore, in theory the d-scores seem to benefit the MSP group in terms of size, but in reality it is most certainly not the case, and have therefore been interpreted for the benefit of the reader as being to the advantage of SARFU.

When the significances are further examined, it can once again be seen that where the MSP has outperformed the SARFU group the significances are very small, ranging from a d-score of 0.04 to 0.35. The variables are the following: body mass, upper arm girth, calf girth and humerus breadth. The variables in which SARFU performed better are the following: triceps, sub-scapular, supra spinal and medial calf skin folds along with femur breadth. The d-score for the femur breadth is an insignificant –0.04, but for the skin folds the range from d-scores 1.26 to 1.8 indicating very high significances.

When the raw data is examined, body mass, humerus and femur breadth can be interpreted as being of very little or even no significance in terms of determining success or failure in the game of rugby. When upper arm and calf girth are examined, where MSP outperformed SARFU the following can be assumed: the d-scores range from small to medium effect, indicating that in reality they have no effect on the game itself.

An interpretation is that MSP outperformed SARFU in these above two variables because of the fact that they had larger skin folds in (in those areas and overall) than SARFU. Excess underlying subcutaneous fat cannot be beneficial for success in rugby, and therefore can be regarded as insignificant.

When the skin fold results are examined, MSP had larger skin folds than SARFU, but as mentioned before, this is not a positive sign. The larger the skin fold, the larger the subcutaneous fat depositions, indicating a lower level of training and conditioning.

This is echoed in the physical motor ability tests, where SARFU outperformed MSP in all the tests.

Therefore, the overall interpretation of this author is that MSP are not at a physical standard, as measured by the anthropometrical tests, to compete with, and attain the level of success that the SARFU group have attained. It is evident that at this current stage of testing, the MSP are far from the level of SARFU.

4.3) DISCRIMINANT ANALYSIS

The process of discriminant analysis has already been fully described in chapter two. The two groups were compared to one another according to their predictive function. The results will be subsequently discussed. Before that discussion however, a look will be taken at the different factors that are the best discriminators.

The STEPDISC-Procedure of the SAS computer programme package (SAS Institute Inc., 1989:1493) was used to obtain the results in Table 4.

Table 4) Indications of best discriminating factors, using stepwise discriminant analysis

NUMBER	VARIABLE	PARTIAL R**2	F STATISTIC
1	Vertical Jump (cm)	0.5498	90.384
2	Tricep SF (mm)	0.2461	23.83
3	Mass (kg)	0.1675	14.491
4	Speed (sec)	0.0921	7.205
5	Speed Endurance (sec)	0.0899	6.918
6	Forward Pass (pts)	0.0756	5.64
7	Pass for Accuracy 7m (pts)	0.1041	7.901
8	Running & Catching (pts)	0.1537	12.169
9	Supra Spinal SF (mm)	0.059	4.142

The above table illustrates those factors have the best discriminating qualities with regards to the SARFU and MSP groups. When the table is examined, the following is clear:

1) Vertical jump is the best discriminator between the two groups. Then, in order of descending significance, you find that tricep skin fold discriminates second best followed by mass, speed, speed endurance, forward pass, pass for accuracy over 7m, running and catching and lastly supra spinal skin fold.

2) When the above factors are examined, the following can be seen:
   - 3 factors are sports-specific, namely forward pass, pass for accuracy 7m and running and catching;
   - 3 factors are motor-physical, namely vertical jump, speed and speed endurance; and
   - 3 factors are anthropometrical, namely tricep skin fold, mass and supra spinal skin fold.

These results are interesting, as it shows an equal distribution of factors that need to be taken into account when discussing talent and the prediction thereof. An overall view needs to be developed and maintained, as it is not really known what determines talent. In past discussions it was stated that all the factors need to be taken into account. The above table and its interpretation prove that point quite clearly.

4.3.1) Comparison of predictive functions of the two groups

The classification functions of the two groups were obtained by using the DISCRIM-Procedure of the SAS (SAS Institute Inc., 1989), using the selected factors from Table 4. The Coefficients for the functions are given in Table 5.

Table 5) Coefficients of classification functions of the SARFU and MSP groups

VARIABLE	GROUP	
	MSP	SARFU
	Less Talented	More Talented
Constant	-49.8491	-65.9543
Mass (kg)	0.54992	0.81921
Supra Spinal SF (mm)	-0.06393	-0.30555
Speed Endurance (sec)	0.57463	-0.08898
Vertical Jump (cm)	0.23838	0.5305
Pass for Accuracy 7m (pts)	0.8413	-0.19199
Speed (sec)	0.12807	-0.21251
Running & Catching (pts)	0.52455	1.51961
Humerus Breadth (cm)	3.67529	2.33513
Forward Pass (pts)	0.87915	1.70534

The classification function for a group is of the form:

$Y = A + B_1X_1 + B_2X_2 + \ldots + B_9X_9$ with A the constant coefficient and B_1 to B_9 the coefficients of the factors X_1 to X_9 as displayed in Table 5.

4.3.2) Cross-validation of the classification function of discriminant analysis

To determine the cross-validity of the two groups, each player's values for the nine factors were substituted into each of the two classification functions to obtain two values. A player was classified into the group corresponding to the largest of the two values (Rao, 1973). Table 6 gives the results of these classifications.

Table 6) Cross-validation of the categorised discriminant analysis

	Into Group: MSP	SARFU	TOTAL
Member from group: MSP	41	0	41
	100%	0	100%
SARFU	0	36	36
	0	100%	100%

When examining above table, it can be seen that both groups have been 100% correctly categorized according to discriminant analysis. Therefore, the method of discriminant analysis is 100% accurate in categorizing or sorting less talented from more talented individuals according to players used in this research.

4.4) LOGISTICAL REGRESSION

The formula that was applied to the results of the two groups, namely MSP and SARFU, to determine the possibility of the discriminating variables can be seen below:

$$P(X) = \frac{\exp (A + B_1X_1 + B_2X_2 + B_3X_3 + B_4X_4)}{1+ \exp (A + B_1X_1 + B_2X_2 + B_3X_3 + B_4X_4)}$$

After the whole process of logistical regression was completed, it was found that, where in the case of discriminant analysis where 9 variables were found to be accurate in predicting success, in logistical regression 4 variables were found to be acceptable variables in the analysis of maximum likelihood estimates. They are the following variables:

1) Calf SF
2) Pass for accuracy 7m
3) Running and catching
4) Speed

When these above variables are further analyzed, the following is found: there is also representation from each of the three categories of factors. Calf SF falls into the anthropometrical category and speed falls into the motor-physical category. Pass for accuracy 7m and running and catching both fall into the sport-specific category however.

Whether this possesses any significance will remain to be seen however. Each of the above variables had certain parameter estimates assigned to them during the course of the investigation. These parameter estimates are listed below:

1) Constant (A)	=	-46.4946
2) Calf SF (B_1)	=	-0.4302
3) Pass for accuracy 7m (B_2)	=	2.0211
4) Running and catching (B_3)	=	-3.0765
5) Speed (B_4)	=	9.7381

When these parameter estimates are substituted into the equation, you get the following model:

$$P(X) = \frac{\exp(-46.4946 - 0.43\,\text{calf SF} + 2.02\,\text{pass for acc. 7m} - 3.08\,\text{running and catching} + 9.74\,\text{speed})}{1 + \exp(-46.4946 - 0.43\,\text{calf SF} + 2.02\,\text{pass for acc. 7m} - 3.08\,\text{running and catching} + 9.74\,\text{speed})}$$

If the probability is high (> 0.5) then the player is categorized into the MSP group. If the probability is low (< 0.5) then the player is categorized into the SARFU group.

When the cross-validation is done on the categorizing into the two respective groups according to the probability function of logistical regression, the following can be seen:

Table 7) Cross-validation of the probability function of logistical regression

	Into Group: MSP	SARFU	TOTAL
From Group: MSP	41	0	41
	100%	0	100%
SARFU	0	35	35
	0	100%	100%

As can be seen from the above table, logistical regression is also successful in 100% categorization of the two groups. Therefore, the method of logistical regression is 100% accurate in categorizing or sorting less talented from more talented individuals.

As stated at the beginning of this study, one of the stated aims was to determine which of the above two methods were more accurate in predicting or discriminating the more from the less successful individuals. As can be seen from the above results, both methods are equally accurate in predicting success.

Since that has been determined, the next stated aim was to determine if you could have a more accurate method if the two methods were combined. As was mentioned previously in this chapter, the discriminant analysis method has nine variables that discriminate better than the rest, whereas logistical regression has four. For recap purposes these factors will be listed below

Table 8) Table listing factors found to be best determinants of success in both predictive function methods

Discriminant Analysis Factor	Match	No Match	Logistical Regression Factor
Pass for Accuracy 7m (pts)	√		Pass for Accuracy 7m
Speed (sec)	√		Speed
Running & Catching (pts)	√		Running & Catching
Supra Spinal SF (mm)		X	Calf SF
Speed Endurance (sec)		X	N/A
Vertical Jump (cm)		X	N/A
Humerus Breadth (cm)		X	N/A
Forward Pass (pts)		X	N/A
Mass (kg)		X	N/A

What can be seen is that three factors or variables correspond in both methods. These are the pass for accuracy 7m, speed and running & catching. Pass for accuracy 7m and running & catching fall into sport-specific category, whereas the speed falls into the motor-physical category. A discussion of the significance of this will follow in chapter five.

A further step was taken, whereby the four factors found to be successful determinants in the logistical regression method were substituted into the discriminant analysis method. The factors were pass for accuracy 7m, calf SF, running & catching and speed. The following was found, and is illustrated in the table below:

Table 9) Table illustrating results obtained when logistical regression factors are substituted into discriminant analysis method

	Into Group: MSP	SARFU	TOTAL
From Group: MSP	40 93%	3 6.98%	43 100%
SARFU	5 13.89%	31 86%	36 100%
TOTAL	45 56.96%	34 43.04%	79 100%

From the above it can be seen that when those four factors are substituted into the discriminant analysis method, results are obtained that are less than 100% prediction accuracy. In the above scenario, there is a 93% prediction accuracy with the MSP group, whereas there is a 86% prediction accuracy where the SARFU group is concerned. The significance of the findings reported in the above table will also once again be discussed in chapter five.