CHAPTER FIVE

HYBRID BARS ANALYSIS

Introduction

This chapter is designated to address three issues pertaining to hybrid BARS, and is divided into five sections. The first section describes the sample of respondents who undertook the survey. This review provides a context for examining subsequent findings. The second section considers the assumption that competencies, identified through the hybrid BARS process, are perceived to be important by a diverse range of soccer stakeholders. The third and fourth sections assesses the classification, by expert panels, of competencies into performance dimensions. Lastly, the fifth section evaluates the internal-consistency of the instrument used in the present study.

OVERVIEW OF RESPONDENTS

Findings in this section provide background information about the survey and respondents. Primarily, discussion focuses upon survey response-rates, and various characteristics unique to each sub-group. In particular, respondent age and soccer background (playing experience and refereeing qualifications) information are reviewed.

Survey Response-Rates

Overall, data were collected from 173 respondents. This response-rate represented 61% of the potential sample. A breakdown of response-rates by sub-group is provided in Table 5.1. Figures from this table show acceptable response-rates across all groups. The highest response-rate was from referees (100%), with the lowest coming from players (51%). Response-rates of 72-80% were recorded for coaches, assistant referees, and referee inspectors. These latter response-rates are impressive, particularly as data from assistant referees and referee inspectors were collected via the post. Typically, response-rates are not high from postal surveys (Cohen & Manion, 1994; Krathwohl, 1993), however, returned responses for the present study were far in excess of those normally associated with this method of data collection.

Table 5.1 Response rates by sub-group

Sub-group	Potential Sample (*n*)	Respondents (*n*)	Response Rate (%)
Players	182	92	51
Coaches	23	18	78
Referees	13	13	100
Assistant Referees	36	26	72
Referee Inspectors	30	24	80
Totals	284	173	61

The distribution of respondents, by state of residence, is provided in Table 5.2. Two points stand out from this table. Firstly, the sample was drawn from all states and territories where Ericsson Cup soccer is played. Moreover, most sub-groups were represented from each state and territory. The only exception is the absence of a coach's responses from Western Australia, and a referee's response from the ACT. However, the absence of responses from the two categories in these regions is of minor significance, as only one coach resides in Western Australia, and no referee from the ACT is on the Ericsson Cup panel.

Table 5.2 Distribution of respondents by state or territory

State	Referees		Assistant Referees		Referee Inspectors		Players		Coaches	
	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*
NSW	38	(5)	27	(7)	33	(8)	32	(29)	39	(7)
Victoria	15	(2)	23	(6)	17	(4)	27	(25)	33	(6)
Queensland	15	(2)	8	(2)	17	(4)	11	(10)	11	(2)
South Australia	15	(2)	19	(5)	8	(2)	13	(12)	11	(2)
Western Australia	15	(2)	15	(4)	8	(2)	9	(8)	0	(0)
ACT	0	(0)	8	(2)	17	(4)	9	(8)	6	(1)
Total		(13)		(26)		(24)		(92)		(18)

The second point that stands out from Table 5.2 was the high proportion of respondents from NSW. Given that NSW has the most Ericsson Cup clubs (n=5), and is the most populous state in Australia (Australian Bureau of Statistics, 1999), the type of distribution presented in Table 5.2 was anticipated.

Age of Respondents

The age-range of respondents for each sub-group is provided in Table 5.3. Not unexpectedly, players presented proportionally greater numbers in the lower age brackets. Conversely, a majority of referee inspectors (83%) were over 50 years of age. Again, this is not surprising given referee inspectors normally assume the inspector role at the conclusion of a refereeing career.

Table 5.3 Age distribution by group

Age (years)	Referees	Assistant Referees	Referee Inspectors	Players	Coaches
	% *n*	% *n*	% *n*	% *n*	% *n*
<21	0 (0)	0 (0)	0 (0)	30 (28)	0 (0)
21-25	15 (2)	19 (5)	0 (0)	37 (34)	0 (0)
26-30	23 (3)	12 (3)	0 (0)	21 (19)	0 (0)
31-35	23 (3)	31 (8)	0 (0)	12 (11)	22 (4)
36-40	38 (5)	19 (5)	4 (1)	0 (0)	33 (6)
41-45	0 (0)	15 (4)	0 (0)	0 (0)	22 (4)
46-50	0 (0)	4 (1)	13 (3)	0 (0)	6 (1)
>50	0 (0)	0 (0)	83 (20)	0 (0)	17 (3)
Totals	(13)	(26)	(24)	(92)	(18)

Given the prime focus of this investigation on refereeing, the age range of referees is worthy of additional comment. In particular, the proportion of referees categorised into the lower-to-mid age ranges shows consistency with referee ages reported in the literature (for example, see Aresu, Bucarelli, & Marongiu, 1979; Johnston & McNaughton, 1994; Koslowsky & Maoz, 1988; Taylor et al., 1990). In the present study, two specific factors determined the age range of referees. Firstly, Soccer Australia incorporates the FIFA policy of not having referees over the age of 45. The adoption of this policy restricts artificially the age-range of Australia's top referees. Secondly, Soccer Australia has implemented a strategy of identifying and recruiting younger referees (early to mid-twenties). The philosophy behind this strategy is to provide a viable career for referees, plus utilise the relatively better fitness levels of younger referees. In combination, these two factors have contributed to keep the age of all referees below 40 years.

Soccer Background

Findings pertinent to the soccer background of respondents reflect two issues, namely, the acquisition of refereeing qualifications by respondents, and the highest competitive level of soccer played by respondents. With respect to the acquisition of refereeing qualifications, a

contrast is evident between respondents involved in officiating-related roles (i.e., referees, assistant referees and referee inspectors), and respondents involved in playing and coaching roles. Figures provided in Table 5.4 highlight this disparity. The contrast was expected, particularly given referees, assistant referees and referee inspectors require formal refereeing qualifications as a pre-requisite to their role. Conversely, players and coaches require no such qualification. Moreover, formal refereeing qualifications for players or coaches present no professional advantage, hence the limited number of qualified referees from within their ranks.

Table 5.4 Prevalence of refereeing qualifications in the sample

	Referees	Assistant Referees	Referee Inspectors	Players	Coaches
	% *n*	% *n*	% *n*	% *n*	% *n*
Yes	100 (13)	100 (26)	100 (24)	2 (2)	11 (2)
No	0 (0)	0 (0)	0 (0)	98 (90)	83 (15)

With respect to playing experience, players and coaches have greater experience at the highest levels of soccer (i.e., national and international competition) than their officiating counterparts (see Table 5.5). Again, this finding was expected. Interestingly, four respondents had not played at the equivalent standard of Ericsson Cup or higher. Closer inspection of the data revealed these individuals to be assistant-coaches. However, all head-coaches had played, at least, in the national league. Moreover, almost half (n=8) had international experience.

Table 5.5 Highest competitive playing level by group

Competitive Level	Referees	Assistant Referees	Referee Inspectors	Players	Coaches
	% *n*	% *n*	% *n*	% *n*	% *n*
Juniors	23 (3)	19 (5)	13 (3)	0 (0)	0 (0)
Amateur	0 (0)	15 (4)	8 (2)	0 (0)	0 (0)
Local Association	46 (6)	35 (9)	54 (13)	0 (0)	17 (3)
State League	23 (3)	31 (8)	13 (3)	0 (0)	6 (1)
Ericsson Cup	0 (0)	0 (0)	0 (0)	50 (46)	28 (5)
International	0 (0)	0 (0)	0 (0)	50 (46)	44 (8)
Missing data	*(1)*		*(3)*		*(1)*

The playing experience of referees, assistant referees, and referee inspectors, present a contrast to players and coaches. For the majority of these sub-groups, the highest competitive playing level attained was the local association level. Particularly, no referee, assistant referee, or referee inspector had played at, or above, Ericsson Cup standard.

Summary

The descriptive findings reported in this section provide an overview of the nature and characteristics of the sample for the present study. All sub-groups presented acceptable response-rates, ranging from 100% for referees to 51% for players. Acceptable numbers of respondents were drawn from all states and territories where Ericsson Cup clubs are based, and were inclusive of all ages.

The varying playing and refereeing characteristics of each sub-group were predictable, particularly given the specialised nature of each sub-group's work. When such specialisation is coupled with the elite performance standards these groups are obliged to meet, the preclusion of high playing experience levels for referees, assistant referees, and referee inspectors, and low incidence of referee qualification for players and coaches, was expected.

THE IDENTIFICATION OF REFEREEING COMPETENCIES

In previous studies utilising hybrid BARS research, researchers assessed the degree of support for the performance dimensions only, primarily via a derivation of frequency analysis techniques (for example, see Anshel, 1995; Anshel et al., 1987; Anshel & Webb, 1991; Moore & Webb, 1995; Webb et al., 1994). The attention directed to analysis of performance dimensions *implied* all competencies, identified through BARS procedures, were also well supported. However, such a conclusion may be tenuous, particularly as the assumption had not been subjected previously to analysis. To avoid making similar assumptions, the present study used the same frequency analysis criteria adopted by earlier hybrid BARS research (for assessing dimension importance) to assess the essential nature of *competencies*. In following this process, competencies were classified into one of three categories, namely, 'must-have,' 'should-have' and 'unimportant.' Relevant findings address the research question:

Do major soccer stakeholders support the importance of refereeing competencies identified through the hybrid BARS process?

Overview of Frequency Analysis Results

Data for this analysis were drawn from the *importance* data set. These data were chosen primarily because competencies identified from the BARS process were seen by expert panels to be essential to elite refereeing. Thus, it is reasonable to assume the global sample of respondents would view the same competencies in a similar manner, i.e., important.

Specific criteria for the classification of each competency was described previously in Chapter 4. However, a brief review of criteria follows: 'must-have' – 90% of sample

responded in the *1* or *2* response category for the competency; 'should-have' – 90% of sample responded in the *1*, *2* or *3* response category for the competency; and, 'unimportant' – competencies which failed to meet 'must-have' or 'should-have' criteria.

Subsequent categorisation of each competency, with respect to these classifications, is provided in Table 5.6 (detailed frequency analysis results are found in Appendix 12). Of the 37 competencies, 16 were considered to be 'must-have' competencies, 17 were considered 'should-have' competencies, and four competencies were considered to be 'unimportant.'

Competencies classified as 'must-have' encompassed a range of refereeing responsibilities. These varied from rule application, to fitness, to concentration, to teamwork with assistant referees. However, two interconnected characteristics emerge which overlay the type of competencies within this classification. Firstly, all competencies related to on-field activities. Secondly, and more pointedly, competencies associated with decision-making were prevalent. Such a finding is not surprising given the on-field requirement, and the stated importance, of decision-making in the literature.

The types of competencies classified into the 'should-have' category where diverse with respect to refereeing responsibilities. Competencies associated with risk management procedures, record keeping, and communication, reflect a marked diversity of refereeing responsibilities within this category. However, an emphasis on communication competencies emerges. Particularly, competencies linked to communication with players, coaches, and assistant referees are classified in the 'should-have' category. More specifically, communication competencies pertaining to the 'should-have' categorisation encompassed a variety of verbal and non verbal techniques, e.g., hand-signals, use of whistle, and report writing. While the classification of such competencies into this category indicates a high level of priority for the competencies, their non-classification into the 'must-have' category was not commensurate with the documented need, within officiating literature, for effective communication skills.

The four competencies classified as 'unimportant' relate specifically to activities undertaken away of the competitive context of the game – although it may be argued, to some degree, that competency 1.3, *understands how the game is played (e.g., tactics and strategy , analyse patterns of play)* is demonstrable in the competitive context. Most obviously, competency 5.3, *engages in post match activities (e.g., talk to coaches, attend post game functions)*, can be considered well-removed from this context. However, notwithstanding such limitations, the performance of the remaining three competencies (i.e., 1.3, 3.8, and 5.2) do, to varying degrees, *impact* on on-field performance. This subtle difference in how competencies are applied may account for the variation between these three competencies, and competency 5.3, evident in frequency analysis results (see Appendix 12). Specifically, when compared

Table 5.6 External validation of elite refereeing competencies

	Competency	Must Have	Should Have	Unimportant
1.1	understands and interprets the Laws of Soccer correctly	✓		
1.2	observes incidents and decides on appropriate action (e.g., fouls, player acting, free kick, booking, play-on)	✓		
1.3	understands how the game is played (e.g., tactics and strategy , analyse patterns of play)			✓
1.4	anticipates players actions and reactions (e.g., retaliation to over physical play)		✓	
1.5	keeps a complete record of the game (e.g., bookings, goals, substitutions)		✓	
1.6	works as a team with assistant referees	✓		
1.7	observes play from the best position	✓		
2.1	applies the Laws of Soccer consistently (within each game and over the season)	✓		
2.2	observes, analyses and correctly interprets incidents	✓		
2.3	reacts quickly and effectively to incidents	✓		
2.4	distinguishes between fair and foul play	✓		
2.5	interprets and discriminates between the severity of fouls	✓		
2.6	encourages attacking play		✓	
2.7	distinguishes between advantage and disadvantage	✓		
2.8	manages conflict (communication with players, use of presence and personality)	✓		
2.9	moves to obtain optimum positions (i.e., place with best view and close enough to react effectively)	✓		
3.1	effectively uses the whistle (e.g., volume, tone, timing, length, player reaction)		✓	
3.2	communicates decisions with clear hand signals		✓	
3.3	undertakes report writing and record keeping (e.g., send off reports, administrative reports)		✓	
3.4	communicates (verbal and non-verbal) with players on and off the field		✓	
3.5	mediates disputes between opposing players		✓	
3.6	communicates confidently with assistant referees		✓	
3.7	communicates appropriately with coaches		✓	
3.8	communicates with referee's inspectors (e.g., post match discussions, self reflection, constructive criticism)			✓
4.1	applies the Laws and sanctions (e.g., free kick, yellow and red cards)	✓		
4.2	displays positive behaviours and attitudes		✓	
4.3	manages disputes between players/coaches and match officials		✓	
4.4	monitors player behaviour	✓		
4.5	manages all aspects of the game (e.g., match control/players)		✓	
5.1	prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared)		✓	
5.2	undertakes mental preparation for the match (e.g., visualisation)			✓
5.3	engages in post match activities (e.g., talk to coaches, attend post game functions)			✓
5.4	manages personal anxiety		✓	
6.1	maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)	✓		
6.2	maintains appropriate levels of personal health (e.g., correct diet, weight control)		✓	
6.3	undertakes appropriate risk management procedures (e.g., know legal responsibilities, state of the pitch)		✓	
6.4	maintains concentration during the game	✓		

with the other three 'unimportant' competencies, competency 5.3 is seen to be the least important competency by a noticeable margin.

However, to judge each competency on the face value of their classification is potentially misleading. This point becomes clearer on inspection of Table 5.7. This table lists all 'should-have' and 'unimportant' competencies with respect to 'should-have' criteria, i.e., percentage of responses in the '*1*', '*2*', or '*3*' response options. Of note are the values listed for competencies 2.6, 3.7, 5.4 ('should-have' competencies), and item 5.2 (an 'unimportant' competency). The three 'should-have' competencies are one percentage point from being *rejected* from this classification category, while competency 5.2 is a similar margin from being *accepted* into the 'should-have' category.

Table 5.7 Classification of 'should-have' and 'unimportant' competencies with respect to 'should-have' criteria

Competency and Classification Category		Cumulative %
Should-have competencies		
1.4	anticipates players actions and reactions (e.g., retaliation to over physical play)	95
1.5	keeps a complete record of the game (e.g., bookings, goals, substitutions)	95
2.6	encourages attacking play	90
3.1	effectively uses the whistle (e.g., volume, tone, timing, length, player reaction)	91
3.2	communicates decisions with clear hand signals	94
3.3	undertakes report writing and record keeping (e.g., send off reports, administrative reports)	92
3.4	communicates (verbal and non-verbal) with players on and off the field	97
3.5	mediates disputes between opposing players	92
3.6	communicates confidently with assistant referees	92
3.7	communicates appropriately with coaches	90
4.2	displays positive behaviours and attitudes	97
4.3	manages disputes between players/coaches and match officials	94
4.5	manages all aspects of the game (e.g., match control/players)	96
5.1	prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared)	91
5.4	manages personal anxiety	90
6.2	maintains appropriate levels of personal health (e.g., correct diet, weight control)	95
6.3	undertakes appropriate risk management procedures (e.g., know legal responsibilities, state of the pitch)	95
Unimportant competencies		
1.3	understands how the game is played (e.g., tactics and strategy, analyse patterns of play)	81
3.8	communicates with referee's inspectors (e.g., post match discussions, self reflection, constructive criticism)	86
5.2	undertakes mental preparation for the match (e.g., visualisation)	89
5.3	engages in post match activities (e.g., talk to coaches, attend post game functions)	74

The results presented in Table 5.7 illustrate the fine line that exists, through frequency analysis methods, for accepting or rejecting a competency as belonging to a specific category. As no statistical substantiation for the classification-criteria was apparent from hybrid BARS literature, the development of classification-criteria appears to be arbitrary.

This presents a tenuous circumstance. However, it is recognised that cut-off limits are necessary in the formulation of classification criteria. In respect of the frequency analysis criteria, an arbitrary 90% does seem initially to be excessive, although it could be argued that the higher the cut-off value, the more rigorous the criteria become.

Summary

Findings presented in this section addressed Research Question 2.1: *Do major soccer stakeholders support the importance of refereeing competencies identified through the hybrid BARS process?* The external validation technique employed by the present study demonstrated hybrid BARS is an appropriate mechanism for identifying important refereeing competencies. Only 11% (n=4) of the competencies were viewed as 'unimportant' despite the classification of items, into this category, being seen as problematic. While it is acknowledged that there are no previous benchmarks for determining the success of the hybrid BARS procedure, it could be argued that, at least subjectively, the results indicate broad acceptance of refereeing competencies identified by the BARS expert panels.

VALIDATION OF PERFORMANCE DIMENSION STRUCTURE – IMPORTANCE DATA

As a consequence of competencies being viewed as important, the classification of competencies into performance dimensions requires scrutiny. As indicated in Chapter 3, the classification process relied heavily on the intuitive judgements of two expert panels. In review, the panels identified 37 competencies, which were classified into six performance dimensions. If such a dimension structure is sustainable, it is expected that an empirical factor structure would be derived from subsequent data analysis.

As competencies were identified and selected on the basis of their essential need to elite soccer refereeing, data for this analysis is most suitably derived from the *importance* data set. Therefore, initial results focus on findings derived from these data. Two statistical techniques are used to assess performance dimension structures, namely, factor analysis and Rasch scaling. Results from these analyses address the research question:

Are the performance dimensions, identified by occupational experts, supported using statistical analyses?

Factor Analysis

All *importance* data were submitted to principle component factor analysis, with OBLIMIN rotation, as described in Chapter 4. Subsequent analysis of all 37 competencies produced a factorial solution. Using the eigenvalue criteria of accepting values greater than one, six factors were extracted in 17 iterations. Table 5.8 displays the initial statistics from this

Table 5.8 Principle component factor analysis – initial statistics for Importance data

Factor	Eigenvalue	Percent of variance	Cumulative percentage
1	14.08848	38.1	38.1
2	3.77537	10.2	48.3
3	1.89228	5.1	53.4
4	1.41349	3.8	57.2
5	1.1662	3.2	60.4
6	1.13882	3.1	63.4

analysis, with particular reference to respective eigenvalues and the proportion of variance accounted for by each of the six factors. In determining the suitability of these results, a number of assumptions which underpin factor analysis were addressed.

The first of these is the Bartlett's Test of Sphericity (BTS). BTS searches for the presence of correlations among variables (Hair et al., 1995), and, in doing so, supplies an index of statistical probability that reflects correlations between at least some variables. For the present data set, the BTS provided an index of 4297.055 (df=172, p<.0001). While the significance of this result should not be understated, the proposition by Hair et al. (1995) that increased sample size may amplify the magnitude of any significant correlations was accepted. Subsequently, the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (MSA) was implemented to quantify the degree of interrelations among variables with respect to sample size. The resulting MSA produced an acceptable coefficient of 0.90782. In combination, the results of BTS and MSA indicate a statistically favourable factor solution.

Given the acceptance of the underlying statistical assumptions, recommendations proposed by Stevens (1996) were observed for determining significant factor loadings for each item. Stevens asserted that significant loadings be assessed with respect to sample size, and, as a consequence, avoid the normal practice of "blindly using ... significant loadings greater than .30" (p. 371). Accordingly, a critical value for significant loadings was calculated at 0.38 (n=180), which is significant at 0.01 for a two-tailed test (Stevens, 1996). Factor loadings for all items are detailed in Table 5.9.

Of note in Table 5.9 is the number of competencies which load significantly onto more than one factor (n=5). Although this circumstance is not uncommon in applied research (Hair et al., 1995), the significant loading of items onto more than one factor may confound the interpretation of the factor structure.

Table 5.9 OBLIMIN rotation pattern matrix[1] – Importance data

Competency	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
1.1	0.17832	0.10069	-0.02811	-0.23180	-0.69476	-0.01555
1.2	0.08952	-0.01117	0.03531	0.08679	-0.76878	0.05237
1.3	0.02276	0.04131	-0.08383	0.84840	0.04895	-0.02477
1.4	0.11106	0.27482	0.40966	0.23876	-0.06269	-0.19425
1.5	0.02646	-0.08995	0.75605	0.07425	-0.10214	0.02966
1.6	0.02490	0.29395	0.11989	0.10417	-0.54745	-0.09178
1.7	0.20574	0.40367	0.53061	-0.12860	0.01546	-0.25139
2.1	0.06918	0.66787	-0.09250	-0.00640	-0.26800	0.04170
2.2	0.04440	0.56096	-0.00092	0.28993	-0.27005	-0.08758
2.3	-0.00797	0.72589	0.08136	0.10285	0.04817	0.09718
2.4	0.15928	0.68840	0.06604	0.06769	-0.10936	-0.12694
2.5	-0.04161	0.71218	0.00607	-0.09877	0.07474	0.16514
2.6	-0.19172	-0.02815	0.24841	0.44076	-0.34839	0.32929
2.7	-0.14173	0.30583	0.35282	0.08994	-0.15448	0.32919
2.8	0.16759	0.61997	0.03065	-0.00185	-0.20652	-0.12468
2.9	-0.11963	0.33108	0.61116	-0.10833	-0.03861	0.00905
3.1	0.39586	-0.11585	0.48060	-0.05260	0.08765	0.33050
3.2	0.57320	0.01385	0.30983	-0.06225	0.09508	0.19694
3.3	0.64191	-0.20656	0.46853	0.11287	-0.05453	-0.03162
3.4	0.78789	0.07402	-0.21249	0.15831	-0.13918	-0.03754
3.5	0.76358	-0.06477	0.12488	0.25019	0.03549	-0.03302
3.6	0.55749	-0.09543	0.24223	-0.02528	-0.23752	0.20615
3.7	0.59264	0.04029	0.05297	0.10971	-0.06385	0.16296
3.8	0.48071	0.12878	0.00248	0.34232	0.25791	0.10680
4.1	0.78245	0.04581	0.05709	-0.10576	-0.05273	-0.01556
4.2	0.72025	0.14770	0.05271	-0.12685	-0.06425	0.13277
4.3	0.69958	0.03428	0.10694	0.16200	0.02453	0.07872
4.4	0.81576	0.02102	-0.16814	-0.01923	-0.23116	0.05797
4.5	0.71901	0.22080	0.01198	-0.10810	0.00632	-0.01579
5.1	0.39599	0.06760	0.14260	-0.07344	0.01799	0.48367
5.2	0.30186	0.15472	0.09358	0.08133	0.09872	0.53998
5.3	0.07538	-0.03456	-0.09410	0.01920	-0.00695	0.67399
5.4	0.28717	0.23212	-0.09803	-0.01558	-0.07280	0.58880
6.1	0.76756	-0.00932	-0.00278	-0.16271	-0.09732	0.19937
6.2	0.50449	-0.05043	0.12177	0.08666	-0.05211	0.39890
6.3	0.48089	0.16042	0.17240	-0.07169	0.12205	0.26746
6.4	0.78906	0.06679	-0.11418	-0.11276	-0.05310	0.03147

[1] All significant loadings are outlined, with the primary significant loading being shaded also. 'Secondary' significant loadings are outlined only.

Conversely, one competency (i.e., 2.7, *distinguishes between advantage and disadvantage*) did not load significantly onto any one factor. Under normal factor analysis procedures, this item would be removed from subsequent analysis. However, Hair et al. (1995) claimed that non-significant loading items should not be rejected automatically, but assessed instead against an item's overall contribution to the research. In the context of the present study, therefore, the non-significant loading of competency 2.7 does not indicate necessarily that the item is unimportant to elite soccer refereeing. To expand on this point, it can be logically argued, from a practical perspective, that a referee's ability to distinguish between fair advantage and disadvantage can be representative of a number of performance areas, e.g., decision-making and technical knowledge. Subsequent removal of a competency, such as this, may result in the rejection of a variable which, for all intents and purposes, is vital to the role of refereeing. For these reasons, competency 2.7 was retained for future analysis.

Notwithstanding the implications of this concern, examination of significant loadings within each factor show the factor structure is interpretable (see Table 5.10). With respect to this structure, a number of pertinent points emerge, and are worthy of discussion. Firstly, five competencies loaded significantly onto more than one factor. For the purpose of completeness, these five competencies are listed within each respective factor, and are italicised where significance is secondary to another factor. Secondly, the interpretation of each factor is based on recommendations provided by Hair et al. (1995), where items with higher loadings are considered more important, and, consequently, have greater influence on the name or label selected to represent each factor. Moreover, the order in which items are listed within each factor is based on the magnitude of each item's loading, i.e., the highest loading item is listed first, the next highest loading item is listed second, and so forth. Thirdly, all six factors are labelled successfully, with the practical significance of each factor supported and verified by the Director of Referees at Soccer Australia (Power, G. 1998, pers.comm., 3 March).

Although this interpretable factor solution appears to demonstrate the success of the hybrid BARS process for classifying occupational competencies into broad performance dimensions, three concerns are evident with the derived factors. The first concern relates to the total percentage of variance accounted for by significant factors (i.e., 63.4% – see Table 5.8). This result is markedly below the minimum 75% recommended by Stoskopf et al. (1992). Moreover, the initial statistics presented in Table 5.8 show Factors 4, 5, and 6 account for relatively small portions of the total variance (6.2%). Although there are no broadly accepted decision rules for significant factors that account for minimal levels of variance, their retention is nonetheless problematic.

Table 5.10 Interpretation of factor solution – Importance data

Factor 1: Game Management and Communication
4.4 communicates clearly with players
6.4 maintains concentration during the game
3.4 communicates (verbal and non-verbal) with players on and off the field
4.1 applies the Laws and sanctions (e.g., free kick, yellow and red cards)
6.1 maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)
3.5 manages disputes between players/coaches and match officials
4.2 displays positive behaviours and attitudes
4.5 manages all aspects of the game (e.g., match control/players)
4.3 manages disputes between players/coaches and match officials
3.3 undertakes report writing and record keeping to maintain a complete record of the game (e.g., send off reports, administrative reports)
3.7 communicates appropriately with coaches
3.2 communicates decisions with clear hand signals
3.6 communicates confidently with assistant referees
6.2 maintains appropriate levels of personal health (e.g., correct diet, weight control)
6.3 undertakes appropriate risk management procedures (e.g., know legal responsibilities, state of the pitch)
3.8 communicates with referee's inspectors (e.g., post match discussions, self reflection, constructive criticism)
5.1 prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit
3.1 effectively uses the whistle (e.g., volume, tone, timing, length, player reaction)

Factor 2: Decision Making
2.3 reacts quickly and effectively to incidents
2.5 interprets and discriminates between the severity of fouls
2.4 distinguishes between fair and foul play
2.1 applies the Laws of Soccer consistently (within each game and over the season)
2.8 manages conflict (communication with players, use of presence and personality)
2.2 observes, analyses and correctly interprets incidents
1.7 observes play from the best position

Factor 3: Referee Techniques
1.5 keeps a complete record of the game (e.g., bookings, goals, substitutions)
2.9 moves to obtain optimum positions (i.e., place with best view and close enough to react effectively)
1.7 observes play from the best position
3.1 effectively uses the whistle (e.g., volume, tone, timing, length, player reaction)
3.3 undertakes report writing and record keeping (e.g., send off reports, administrative reports)
1.4 anticipates players actions and reactions (e.g., retaliation to over physical play)

Factor 4: Game Knowledge
1.3 understands how the game is played (e.g., tactics and strategy, analyse patterns of play)
2.6 encourages attacking play

Factor 5: Rule Interpretation
1.2 observes incidents and decides on appropriate action (e.g., fouls, player acting, free kick, booking, play-on)
1.1 understands and interprets the Laws of Soccer correctly
1.6 works as a team with assistant referees

Factor 6: Personal Attributes
5.3 engages in post match activities (e.g., talk to coaches, attend post game functions)
5.4 manages personal anxiety
5.2 undertakes mental preparation for the match (e.g., visualisation)
5.1 prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared)
6.2 maintains appropriate levels of personal health (e.g., correct diet, weight control)

Secondly, concerns exist with the item composition of various factors. As shown in Table 5.10, Factor 4 contains two items only. Occurrence such as this presents interpretation problems, as a factor defined by a small number of items comes very close to the factor being variable specific (Stevens, 1996). The problem for the factor structure is further compounded, and contrasted, by the 18 items (almost half the total number of items) which loaded significantly onto Factor 1. In combination, the composition of Factors 1 and 4 provide examples of inconsistencies that impinge on the integrity of the factor structure.

Thirdly, an overt match between the factor interpretation and the performance dimension model (developed through the hybrid BARS procedure) is not distinct. Although both the dimension and factor analysis models generated six underlying constructs, their item composition is not identical. Notably, Factor 1 comprises 16 items drawn from three of the original performance dimensions. The trend of a factor encompassing items located originally within a number of dimensions is typical of the factor solution, with no exceptions to this observation in any factor.

Given these three concerns, it is prudent to subject the data to further analysis. Two additional factor analysis procedures were employed for this purpose. Firstly, scree-test criteria were applied to the statistics generated from the initial factor solution, and, secondly, additional principle components analysis, comprising random split half samples, were administered.

Scree-Test Criterion

To satisfy the scree-test criterion, eigenvalues are plotted against factors in order of extraction. The point at which the curve begins to flatten is considered to be the acceptable maximum number of factors to extract (Tabachnick & Fidell, 1996). This process identifies the number of factors that can be extracted before unique variance begins to dominate the common variance structure (Hair et al., 1995). Scree-test results for the *importance* data set are shown in Figure 5.1, and are based on output which emanated from the original principle components analysis (see Table 5.8).

As shown in Figure 5.1, the curve begins to flatten noticeably at Factor 3. Thus three extracted factors only can be considered significant. This outcome was unexpected given that, in general, scree tests result in *more* factors being considered significant than the eigenvalue criteria (Hair et al., 1995). The three significant factors extracted through the use of scree test criteria is inconsistent with this observation and the six factors extracted initially using the eigenvalue criteria.

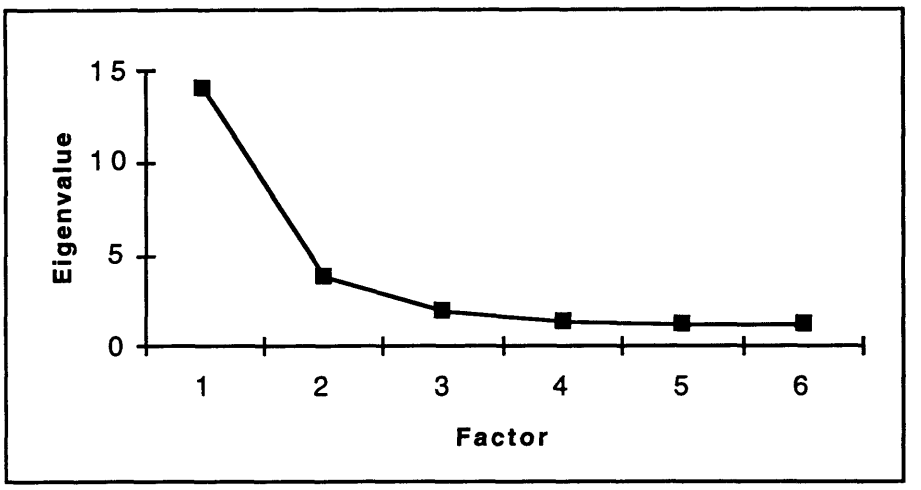

Figure 5.1 Eigenvalue plot for scree-test criterion – Importance data

To minimise the chance that results reported to this point arose from possible statistical anomalies, the data were subjected to a series of random split-half analyses using principle components factor analysis. Findings from this procedure, if similar to those presented thus far, provide additional evidence that the factor structure may be questionable and/or unstable.

Split-half Analysis

Using the criteria of eigenvalues greater than one, described previously, six randomly selected split-half samples from the *importance* data set were submitted for additional principle components factor analysis (with a maximum 50 iterations). Table 5.11 provides a summary of the initial statistics resulting from these tests. Of the six tests, three failed to converge in the allocated 50 interations. Where convergence did occur, the number of extracted factors, and accountable cumulative percentage of variance, was inconsistent. This result demonstrates that factor instability is evident.

Table 5.11 Overview of initial statistics results for Split-half tests – Importance data

Split-half Test	Extracted Factors	Percent of Variance	Convergence
1	-	-	Failed to converge
2	6	69.7	47 iterations
3	6	76.0	22 iterations
4	-	-	Failed to converge
5	10	71.6	37 iterations
6	-	-	Failed to converge

The discrepancies within these split-half factor solutions was typified in results for Test 5 (see Appendix 13). This test provides the most extreme number of extracted factors (n=10),

which is not only more exaggerated than results produced by the original factor solution (and other split-half results), but the interpretation of factors is again problematic.

In particular, the composition of factors in Test 5 is inconsistent, despite some factors being defined by the same (or similar) name as the initial factor solution. Moreover, only five factors could be interpreted with any clarity. Three confounding problems are observed in the remaining five factors, and these are:

1. items which are so disparate in their respective meanings, logical factor labels can not be determined;

2. factors containing items which, for applied settings, are difficult to label; and,

3. factors with a small number of significantly loading items. Specifically, three factors contain two or less items. As a consequence, only one of these factors can be interpreted, the other two factors remained undefined.

From the range of factor analysis findings provided thus far, it seems that the hybrid BARS classification of competencies into performance dimensions is not as stable as first presumed. However, it has been shown that ordinal variables, as well as highly correlated factors, confuse the factor analysis procedure (Schumacker & Linacre, 1996). For these reasons, it is recommended that if factor analysis produces factors that are difficult to interpret, results should be confirmed by Rasch analysis (Bond, 1994).

Rasch Analysis

As noted in Chapter 4, the Rasch analysis technique assumes unidimensionality of data. If data fit the model, unidimensional data are implied. Alternatively, if it is shown that data do not fit the model, multidimensional data are implied, thus raising the possibility that latent traits (analogous to performance dimension) are present in the data.

Consequently, all 37 competencies from the *importance* data set were submitted for Rasch scaling. The fit-statistics provided by this analysis are given in Table 5.12.

Table 5.12 Rasch fit-statistics for Importance data

Criteria	Statistic
Item Estimates	
infit mean square	0.99
infit *t*	-0.22
reliability of estimate	0.88
Case Estimates	
infit mean square	1.07
infit *t*	0.00
reliability of estimate	0.89
Item Consistency Index	0.93

Statistics pertinent to this discussion are found within the item estimates and item consistency index (additional statistics provided in Table 5.12 are discussed later in this chapter, and in Chapter 6). As evidenced by Table 5.12, the fit statistics are acceptable (infit mean square close to 1, and infit *t* close to 0). Thus, the fit-statistics suggest a unidimensional data set.

The item fit-map, produced in conjunction with this analysis, provides additional verification of this finding (see Figure 5.2). If data were to be indicative of underlying traits or constructs, asterisks on the item map would be well spread, with the majority of items lying outside the acceptable fit limits. Of the 37 competencies submitted to scaling, eight fell outside the fit limits. With respect to these outlying items, only one competency (i.e., competency 5.3) falls outside the fit limits by a notable magnitude. The location of this competency, on the item fit-map, is not surprising given the low acceptance of this competency reported formerly in the frequency analysis results.

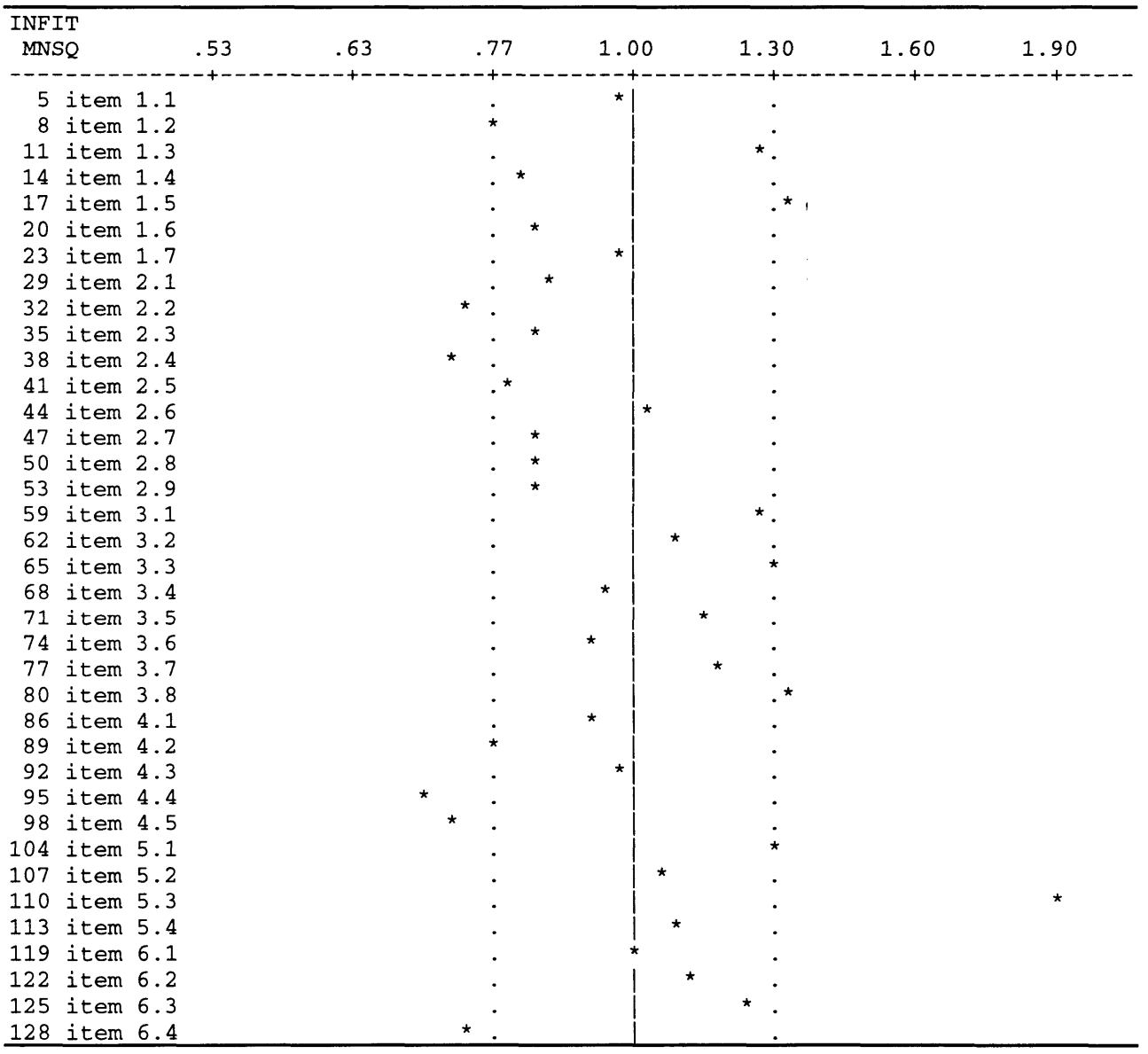

Figure 5.2 Item fit-map for Importance data

The final indication of unidimensional data displayed in Table 5.12 is the *item consistency index*. This index measures the extent to which items are homogeneous. The internal consistency index provided by the Quest software for polychotomously scored data are analogous to Cronbach's alpha (Adams & Khoo, 1996). With a suggested acceptable alpha level of 0.75 (Ellis, 1994), the item consistency index of 0.93 for this scale suggests items were aligned conceptually and indicative of one over-riding perspective.

Summary

The hybrid BARS methodology enables occupational competencies to be identified and, subsequently, grouped into broad performance dimensions. In this study, 37 competencies were identified, and these were grouped into six dimensions. Of particular interest was whether similar performance dimensions could be substantiated through a statistical analysis of the data. This was the focus of this section, and addressed Research Question 2.2: *Are the performance dimensions, identified by occupational experts, supported using statistical analyses?*

In addressing Research Question 2.2, both principle components factor analysis and Rasch scaling were employed. From these analyses, some confounding issues arose. For example, the factor analysis procedure produced six factors, but these factors did not mirror the six competency groupings identified by the BARS expert panels. Further, the factors displayed low levels of accountable variance, and lacked stability.

When Rasch scaling techniques were applied to the data, results suggested that there was a single trait being measured, as only one competency displayed a notable misfit to the Rasch model. The implication is that the data should be considered as a single construct, and other groupings are possible within that construct due to the close relationships among different competencies.

When the factor analysis and Rasch scaling results are viewed together, a clearer picture emerges. Rasch scaling lends support, and provides reason, for the lack of stability identified by the factor analysis. As a result, it can be concluded that, in this study, the performance dimensions are not unique, and cannot be supported through a statistical analysis of the *importance* data.

VALIDATION OF PERFORMANCE DIMENSION STRUCTURE – PREPAREDNESS AND IMPROVEMENT-PRIORITY DATA

In so far as statistically sustainable factor structures were not found within the *importance* data, possible factor structures within the *preparedness* and *improvement-priority* data sets are investigated in this section to ensure completeness of analysis. In doing so, the analysis explores the possibility that respondents had some other form of subliminal factor structure

in mind when assessing each item. The implication is that respondents may have overlooked the dimension structure grounded originally in the *importance* data. In doing so, this section addresses also the research question:

Are performance dimensions, identified by occupational experts, supported using statistical analyses?

Analysis of data in this section employs the same techniques used to analyse the *importance* data, i.e., factor analysis (eigenvalue criteria, scree-test, and split-half tests) and Rasch scaling (fit statistics and internal consistency index).

Factor Analysis

In accordance with procedures undertaken for analysing *importance* data, the *preparedness* and *improvement-priority* data sets were submitted to principle component factor analysis with OBLIMIN rotation. Initial statistics and associated assumptions from these analyses indicates data to be factorable. Table 5.13 provides a summary of these results.

Table 5.13 Initial statistics for Preparedness and Improvement-priority data

	Preparedness data	Improvement-priority data
Factors extracted (eigenvalue ≥ 1)	6	7
Accountable variance ()	65.6	67.3
Bartlett's chi-square	4373.9180[1]	4415.8605[1]
Kaiser-Meyer-Olkin measure of sampling adequacy	0.92740	0.90827
Iterations	31	34

1. (df=172, p<.001)

Subsequent pattern matrices for both data sets present confounding results (see Tables 5.14 and 5.15). With respect to the *preparedness* data, three issues arise. Firstly, Factor 5 contains 18 items with significant loadings (i.e., ≥0.38). This represents almost half of the total number of items loading onto one factor. The high number of significantly loading items onto one factor was observed previously in the *importance* data set. Secondly, in association with the first issue, is the small percentage of total variance (3.3%) accounted for by Factor 5 items. In combination, these two issues show a factor, with a large number of significantly loading variables, accounting for relatively small amounts of explained variance.

A third complicating issue is competency 4.2 not loading significantly onto any one factor, while three competencies loaded significantly onto more than one factor. While this circumstance is not uncommon, in combination with the previously noted issues, these

Table 5.14 OBLIMIN rotation pattern matrix – Preparedness data

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
1.1	-0.16965	0.15493	0.33388	-0.02762	0.73419	-0.01760
1.2	0.00640	-0.08505	0.21311	0.11291	0.76671	0.02711
1.3	-0.00377	-0.09257	0.18171	0.09379	0.20584	0.57708
1.4	0.14076	0.14748	0.07842	-0.17394	0.60628	0.18605
1.5	0.13834	0.13804	0.71384	-0.03246	0.09147	0.07635
1.6	-0.22506	0.10769	0.05135	0.09468	0.77354	-0.03064
1.7	0.11341	0.16829	0.24342	-0.28055	0.57840	0.12134
2.1	0.01473	0.01681	-0.11380	0.02832	0.80897	0.02792
2.2	0.01597	-0.00678	-0.09151	-0.00802	0.81538	0.14954
2.3	0.23088	-0.03362	0.21948	-0.19078	0.60101	0.20890
2.4	0.20798	-0.12515	0.00397	0.12883	0.72255	0.05373
2.5	0.16310	-0.25795	-0.05423	0.20713	0.71098	-0.07525
2.6	-0.09953	0.10027	-0.10562	0.18403	0.39902	0.55716
2.7	-0.03323	0.12504	-0.11098	0.04597	0.69492	0.20064
2.8	0.22571	0.00943	-0.14423	0.02716	0.56770	0.09914
2.9	0.26227	0.00118	-0.10030	0.00212	0.50499	0.17496
3.1	0.19704	0.46317	0.26449	0.28576	-0.19506	0.10817
3.2	0.08359	0.47396	0.26391	0.37551	-0.00581	0.01427
3.3	0.17167	0.05491	0.51025	0.53537	-0.11215	0.08459
3.4	0.56127	0.13633	-0.24489	0.13984	0.28526	-0.03361
3.5	0.62238	0.10809	-0.01693	0.21504	-0.07562	0.18761
3.6	0.52976	0.34369	0.05341	-0.00397	0.14689	-0.08900
3.7	0.73304	-0.02738	0.27954	0.02881	-0.00742	0.06143
3.8	0.18649	-0.09100	-0.00118	0.70161	0.01864	0.02485
4.1	0.18317	0.34646	-0.01104	0.25669	0.41017	-0.19328
4.2	0.35257	0.27079	0.02162	0.26772	0.35212	-0.19706
4.3	0.61236	0.16105	0.12351	0.16149	0.02540	0.06516
4.4	0.52009	0.07777	-0.02600	0.16716	0.43409	-0.08860
4.5	0.36562	0.25371	0.03397	0.15422	0.40714	-0.22314
5.1	-0.04309	0.82064	0.09367	0.10997	-0.06758	-0.02450
5.2	-0.12603	0.20992	-0.17693	0.67771	0.25008	0.09234
5.3	0.28913	0.43499	-0.32710	0.04830	-0.24664	0.35299
5.4	0.36340	0.54047	-0.20440	-0.11078	0.07749	0.12515
6.1	0.03592	0.82559	0.02770	-0.02937	0.15059	-0.05977
6.2	0.02253	0.68659	0.06635	0.18493	-0.01924	0.14110
6.3	-0.00900	0.18503	0.05791	0.65719	-0.00903	0.10289
6.4	0.35151	0.34380	-0.00186	0.07257	0.41020	-0.13319

Table 5.15 OBLIMIN rotation pattern matrix – Improvement-priority data

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
1.1	0.05621	-0.08571	0.14807	0.10739	0.03831	0.77614	-0.07931
1.2	0.01999	0.04026	-0.09080	0.09291	0.13721	0.71114	-0.11313
1.3	-0.07259	-0.05789	-0.03532	0.73384	-0.05890	0.31037	-0.01949
1.4	-0.02221	0.13045	0.02529	0.08091	-0.07169	0.59325	0.03249
1.5	0.06995	0.08414	0.60027	-0.13625	0.07327	0.29681	-0.07750
1.6	-0.26944	0.50007	0.13647	-0.00186	0.13467	0.28900	-0.13212
1.7	0.20944	0.33239	0.08947	-0.23092	0.10924	0.44259	-0.05216
2.1	-0.01822	0.59131	-0.06248	0.17407	0.08232	0.18065	-0.02135
2.2	0.09913	0.57820	-0.05384	0.48019	-0.00553	0.00235	0.06079
2.3	0.47385	0.54297	0.06831	-0.07704	-0.09491	0.07046	0.13718
2.4	0.07226	0.59794	-0.08912	0.05207	0.06063	0.24144	-0.07067
2.5	-0.01835	0.48830	0.00194	0.03206	-0.01515	0.28987	0.09190
2.6	-0.17362	0.38049	0.45254	0.19684	0.17482	0.06222	-0.02685
2.7	0.07661	0.75950	0.11402	0.04186	0.04296	0.00478	-0.02968
2.8	-0.06783	0.58438	-0.29120	0.00949	0.03021	-0.06796	-0.42356
2.9	0.17521	0.64525	0.21972	-0.22594	-0.07821	0.01760	-0.20655
3.1	0.35651	-0.06432	0.45615	0.04794	0.03104	0.05137	-0.29037
3.2	0.49063	0.00840	0.35916	0.01714	-0.03566	-0.00614	-0.38740
3.3	0.23211	0.11248	0.43122	0.32244	0.07394	-0.14141	-0.31340
3.4	0.09240	0.14268	-0.20902	0.06743	0.15254	-0.00584	-0.62679
3.5	0.01765	-0.07491	0.14512	0.13264	-0.09899	-0.01252	-0.85048
3.6	0.12978	0.01233	0.18019	-0.13890	0.23092	0.10604	-0.55824
3.7	-0.05270	-0.02260	0.07669	-0.12219	0.08602	0.13590	-0.78520
3.8	-0.18644	0.26116	0.19800	0.24627	0.39886	-0.20041	-0.37673
4.1	0.51890	0.14246	-0.06996	0.08407	0.07225	0.18316	-0.25027
4.2	0.40826	0.13135	-0.00515	0.12164	0.10614	0.17356	-0.37887
4.3	0.29862	-0.05337	0.10891	0.15498	0.05666	0.07999	-0.54640
4.4	0.26930	0.17260	-0.23931	0.17859	0.07355	0.16759	-0.44103
4.5	0.11236	0.20314	-0.35490	0.07567	0.10330	0.18043	-0.47141
5.1	0.51688	-0.11577	0.27033	0.06444	0.31466	0.12216	-0.02192
5.2	0.21264	0.13119	0.28165	0.25743	0.51971	-0.06290	0.03021
5.3	-0.05562	-0.17285	0.04704	-0.11070	0.90189	0.10751	-0.01268
5.4	0.08022	0.04332	-0.13099	-0.00742	0.91502	-0.05227	0.06833
6.1	0.68634	0.10536	-0.02566	-0.10209	0.21597	0.05140	-0.10732
6.2	0.51066	0.02386	0.08619	0.07609	0.35510	-0.01522	-0.09530
6.3	0.45654	-0.01840	0.12986	0.47910	0.17037	-0.14557	-0.14733
6.4	0.45352	0.07211	-0.17765	0.02642	0.20843	0.26149	-0.22481

findings cast doubt over the adequacy of the factor solution. Moreover, concerns with the factor solution are symptomatic of findings obtained from the *importance* data set.

Results from the *improvement-priority* data show a disbursement of items which is dissimilar to previous data sets. As the pattern matrix displayed in Table 5.15 shows, all factors contained significant loadings. Moreover, items do not congregate predominantly around one factor. This is in marked contrast to results reported for previous data sets. However, a number of competencies (n=5) loaded significantly onto more than one factor, which is consistent with results for the *importance* and *improvement-priority* data.

The interpretation of factor solutions for both data sets is indistinct. For *preparedness* data (see Appendix 14), only four of the six factors can be named. Specifically, the first undefined factor (Factor 2), contains items which are indicative of a wide variety of refereeing responsibilities, including: refereeing mechanics; communication; physical and psychological health; and, post-game social interaction. The second undefined factor (Factor 4), encompasses competencies applicable to record keeping, communication, mental preparation, and risk management. The disparity of competencies within these two factors makes logical labelling difficult. Additionally, the small number of significant loading items for Factor 3 (n=1) and Factor 6 (n=2) pose further threats to the integrity of the factor solution.

From analysis of the *improvement-priority* data, four of the seven factors can be defined clearly and titled (see Appendix 15). Of the undefined factors, Factor 4 is most problematic due to the small number of significantly loading items (n=3). The two other undefined factors (Factors 2 and 3) contain an acceptable number of items, but the disparate relationship between these items make the judicious labelling of the factors difficult.

In summary, the unclear factor structure and low number of items per factor is symptomatic of the factor structure reported previously for the *importance* and *preparedness* data sets. Unequivocal or obvious theoretical (and/or practical) links between items within their respective factors were not apparent for either the *preparedness* or *improvement-priority* data sets.

Subsequent scree-tests conducted on both data sets, using the standardised procedures established for analysing the *importance* data, confirm the inherent inconsistencies for each respective factor structure (see Figure 5.3 and Figure 5.4). For the *preparedness* data, three significant factors can be extracted with any degree of confidence. This is in contrast to the six significant factors extracted using the eigenvalue criteria. Similarly, the scree test on the *improvement-priority* data also extracted three factors, as opposed to seven significant factors using the eigenvalue criteria.

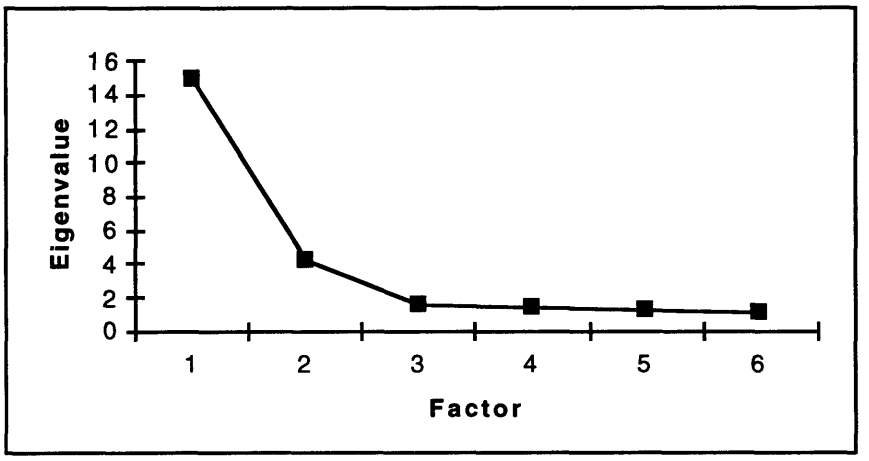

Figure 5.3 Eigenvalue plot for scree test criterion – Preparedness data

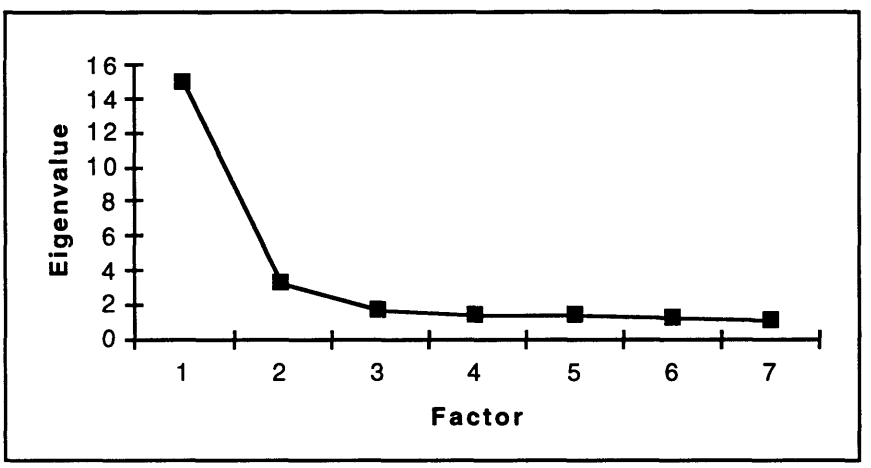

Figure 5.4 Eigenvalue plot for scree test criterion – Improvement-priority data

The ambiguous results provided by the eigenvalue and scree-test criteria for the *preparedness* and *improvement-priority* data are further evidenced in subsequent split-half tests. Consistent with the *importance* data set, six random split-half tests conducted on these data confirmed unstable factor structures. As shown in Table 5.16, not all tests converged in 50 iterations. Of those tests that did converge, the number of extracted factors, and the percentage of accountable variance, is variable with respect to each data set. Again, this variability is consistent with the unstable factor structures derived from the *importance* data split-half tests.

Table 5.16 Split-half tests for Preparedness and Improvement-priority data

Data Source	Split-half Test	Extracted Factors	Percent of Variance	Convergence (iterations)
Preparedness	1	8	70.6	31
	2	7	72.0	33
	3	-	-	Failed
	4	9	69.8	40
	5	-	-	Failed
	6	6	67.4	30
Improvement-priority	1	-	-	Failed
	2	-	-	Failed
	3	7	70.5	43
	4	7	71.3	40
	5	7	69.5	49
	6	9	72.7	40

Given the lack of clarity provided by these factor analysis procedures, Rasch analysis was used to determine the degree of homogeneity in the data sets. The use of this additional statistical procedure is equivalent to the analyses used for the *importance* data set.

Rasch Analysis

Data from the *preparedness* and *improvement-priority* perspectives were submitted to separate Rasch analyses. Subsequent fit-statistics are outlined in Table 5.17 and Table 5.18. These tables show that both data sets conform to the Rasch model, thus indicating unidimensional data.

Table 5.17 Fit statistics for Preparedness data

Criteria	Statistic
Item estimates	
infit mean square	1.00
infit *t*	-0.12
reliability of estimate	0.88
Case estimates	
infit mean square	1.03
infit *t*	-0.34
reliability of estimate	0.96
Item Consistency Index	0.95

Table 5.18 Fit statistics for Improvement-priority data

Criteria	Statistic
Item estimates	
infit mean square	1.00
infit *t*	-0.03
reliability of estimate	0.89
Case estimates	
infit mean square	1.07
infit *t*	-0.18
reliability of estimate	0.95
Item Consistency Index	0.96

Moreover, respective item fit-maps derived from these analyses (see Appendix 16) show that, although eight items are located outside the acceptable fit limits for *preparedness* data – and seven items are located outside these limits for the *improvement-priority* data – in each case four items only were markedly outside the fit limits. The unidimensional nature of both data sets is supported further by each item consistency index being above the acceptable 0.75 value. Accordingly, the Rasch analysis confirms the lack of stable underling constructs in the data, and provide evidence for explaining the poor factor structures within the *preparedness* and *improvement-priority* data.

Summary

Results presented in this section, derived from the *preparedness* and *improvement-priority* data, address Research Question 2.2: *Are the performance dimensions, identified by occupational experts, supported using statistical analyses?* Application of principle components factor analysis techniques to both data sets did not confirm the existence of stable underlying dimensions. This was reflected in the results provided by Rasch scaling, which illustrated that the *preparedness* and *improvement-priority* data sets were essentially unidimensional.

Given these findings, which mirror those reported in the previous section for the *importance* data, the statistical classification of competencies into performance dimensions for the *preparedness* and *improvement-priority* data was not possible. This result implies that the grouping of competencies, by expert panels, is not necessarily unique. More specifically, the results lend evidence to the possibility that the inter-relationships that exist between competencies prevent the identification of sustainable dimension structures, irrespective of which competencies are grouped together.

As a consequence of these findings, it is conceivable that officiating work can not be sub-divided statistically into stable groupings of competencies, and that the refereeing competencies reflect one single soccer 'construct' pertaining to elite soccer refereeing in

Australia. However, the results do not negate the value in identifying and using a set of broad competency groupings, based on expert opinion, for practical purposes.

RELIABILITY OF THE RESEARCH INSTRUMENT

Following rejection of the performance dimension structure, the internal consistency of the instrument was assessed through the analysis of responses to the 37 competencies indicative of effective refereeing performance. Such analysis addresses reliability criteria over-and-above that discussed previously in Chapter 4. Particularly, two statistical procedures were used. Firstly, Cronbach's *alpha* was computed to determine the consistency of all responses to all items within each perspective, and, secondly, a Pearson's Product Moment Correlation coefficient was found to assess the consistency of responses across the perspectives of *importance*, *preparedness*, and *improvement-priority*. The results of this analysis provide evidence to answer the research question: *Do hybrid BARS instruments demonstrate acceptable reliability characteristics?*

Internal Consistency Across All Data

The previous section detailed the lack of obvious dimensions, assessed using factor analysis and Rasch scaling, in the *importance*, *preparedness*, and *improvement-priority* data sets. Particularly, Rasch analysis showed that the data sets were unidimensional, and, thus, measured a single perspective. This result indicated subjects responded to items independently of the performance dimension structure. However, it needs to be established if respondents were responding to all items, within each perspective, with one single belief or knowledge construct in mind.

In this regard, confirmation of data homogeneity can be obtained through computing measures of internal consistency, such as Cronbach's *alpha* coefficient. This measure is most suitable to the present study for two reasons. Firstly, only one administration of the instrument is required (Wiersma, 1991), and, secondly, the technique accounts for polychotomously scored items such as those provided using Likert scales (Crowl, 1993). As a derivation of the Kuder-Richardson formula, the effect of Cronbach's *alpha* is to produce a coefficient that is equivalent to the average of all possible split-half coefficients (Burns, 1997). In doing so, it measures the consistency of responses on each item. Item analysis procedures built into the Quest software for Rasch analysis provide this detail in the form of the *item consistency index* (Adams & Khoo, 1996).

Item indexes for each perspective have been reported previously in discussion centred around possible dimension or factor structures within each perspective. However, for purposes of clarity, the item consistency indexes for each perspective are listed in Table 5.19.

Table 5.19 Item consistency indexes

Perspective	Item Consistency Index
Importance	0.93
Preparedness	0.95
Improvement-priority	0.96

The measures of internal consistency for these data sets are all greater than 0.90. Given the widely accepted decision rule for acceptable reliability coefficients, i.e., > 0.75 (see Ellis, 1994), these Cronbach's *alpha* coefficient equivalents are favourably high. In practical terms, the result indicated that a single attitude, or belief, is being expressed about the nature of elite soccer refereeing through the perspectives of *importance*, *preparedness*, and *improvement-priority*.

Consistency Across Perspectives

In assessing the consistency of responses across each perspective, the analysis undertaken (and subsequent conclusions) present a second measure of internal consistency. As the perspectives of *importance*, *preparedness*, and *improvement-priority* represent patterns of thought, they can be measured for accuracy through correlation coefficients.

Before proceeding with this analysis, it is important to revisit the conceptual relationship between perspectives as described in Chapter 4, i.e.,

Importance – Preparedness = Improvement-priority

Therefore, in the context of obtaining a correlation coefficient, when item scores for *preparedness* are subtracted from scores for *importance*, the resulting scores should correlate highly with the *improvement-priority* scores.

For deriving this measure, all item responses from the three perspectives where subjected to Rasch analysis on one single scale. The approach to scaling was required to avoid the possibility of untrue item estimates being derived if estimates were calculated on three separate scales (i.e., as scales are based on a floating mean of zero, three separate scales would not give a true indication of the differences between perspectives).

Fit statistics showing an infit mean square of 0.99 (sd=0.21) and an infit t of -0.28, indicated an acceptable fit of model to data (see Table 5.20). The item separation index of 0.96 showed that items were well separated and their scale positions relatively stable.

Table 5.20 Fit statistics for all data

Criteria	Statistic
Item Estimates	
infit mean square	0.99
infit *t*	-0.28
reliability of estimate	0.96
Case Estimates	
infit mean square	0.98
infit *t*	-0.68
reliability of estimate	0.96
Item Consistency Index	0.96

Item estimates generated by the Rasch analysis provided the interval level scores required by Pearson's correlation, and are detailed in Table 5.21. This table also provides scores calculated for the difference between *importance* and *preparedness* scores, in the 'Difference' column. Subsequently, scores from this column were correlated with item estimates listed in the 'Improvement-priority' column using Pearson's Correlation Coefficient (r).

For purposes of analysis, Pearson's Correlation Coefficient (r) is given by the equation:

$$r=\frac{N\sum XY-(\sum X)(\sum Y)}{\sqrt{\left[N\sum X^2-(\sum X)^2\right]\left[N\sum Y^2-(\sum Y)^2\right]}}$$

where X = (*importance – preparedness*) and Y = *improvement-priority*.

Using the SPSS statistical package, a correlation coefficient was calculated at 0.93 (significant at $p \leq 0.001$). This coefficient accounted for 86% of the total variance. The net result of this finding is that an underlying pattern of responses emerged which is consistent with the intended meanings of each perspective, i.e, if an item is designated a high value on *importance*, yet a low value for *preparation,* it will stand that the same item be given a high value for *improvement-priority.* Other permutations of the expected pattern included:

- low importance and low preparation = low improvement;
- low importance and high preparation = low improvement; and,
- mid importance and mid preparation = mid improvement.

Table 5.21 Item estimates with respect to perspective

Item	Importance	Prepared	Improvement-priority	Difference
1.1	1.88	-0.21	0.01	2.09
1.2	1.37	-0.46	0.18	1.83
1.3	-0.13	-0.95	-0.4	0.82
1.4	0.51	-0.57	-0.04	1.08
1.5	0.69	0.29	-0.68	0.40
1.6	1.03	-0.49	0.13	1.52
1.7	0.89	-0.33	0.12	1.22
2.1	1.34	-0.67	0.28	2.01
2.2	1.18	-0.52	0.17	1.70
2.3	0.87	-0.33	-0.01	1.20
2.4	1.09	-0.39	0.06	1.48
2.5	0.97	-0.52	0.18	1.49
2.6	0.28	-0.56	-0.32	0.84
2.7	0.92	-0.49	-0.01	1.41
2.8	1.05	-0.67	0.08	1.72
2.9	0.86	-0.29	0.02	1.15
3.1	0.05	-0.15	-0.70	0.20
3.2	0.19	-0.11	-0.66	0.30
3.3	0.29	-0.07	-0.70	0.36
3.4	0.70	-0.66	0.05	1.36
3.5	0.16	-0.61	-0.29	0.77
3.6	0.42	-0.15	-0.32	0.57
3.7	-0.03	-0.63	-0.42	0.60
3.8	0.05	-0.56	-0.51	0.61
4.1	1.11	-0.38	-0.06	1.49
4.2	0.65	-0.34	-0.24	0.99
4.3	0.28	-0.57	-0.28	0.85
4.4	0.88	-0.58	-0.03	1.46
4.5	0.80	-0.43	0.01	1.23
5.1	0.28	0.09	-0.78	0.19
5.2	0.06	-0.38	-0.61	0.44
5.3	-0.58	-0.69	-0.76	0.11
5.4	0.07	-0.56	-0.36	0.63
6.1	1.06	-0.06	-0.32	1.12
6.2	0.33	-0.32	-0.41	0.65
6.3	0.31	-0.22	-0.54	0.53
6.4	1.16	-0.25	-0.06	1.41

The correlation coefficient quoted previously illustrates that these expected trends were consistent in the data, and that the internal consistency of the instrument (with respect to the three perspectives) was verified.

Summary

As the perspectives of *importance*, *preparedness* and *improvement-priority*, with respect to each competency, had not been investigated in previous hybrid BARS studies, the internal consistency of responses was open to question. However, this reliability issue was addressed in this section, which used two forms of analyses to answer Research Question 2.3: *Do hybrid BARS instruments demonstrate acceptable reliability characteristics?*

The first of these was Rasch analysis, which generated item consistency indexes (analogous to Cronbach's *alpha*). These were computed to measure the consistency of responses within each perspective. The resulting high indexes were indicative of acceptable levels of internal consistency.

Secondly, the responses to each perspective were assessed through a combination of Rasch analysis and Pearson's Correlation Coefficient to assess the internal reliability of the instrument with respect to each perspective. Item estimates derived from the *improvement-priority* data were correlated with the 'gap' between *importance* and *preparedness* item estimates. The resulting high correlation between the 'gap' and *improvement-priority* estimates provided evidence that subjects were responding to each perspective as intended. In combination, the results presented in this section support the reliability of the research instrument used in this study.

CONCLUSION

This chapter set out to answer three specific research questions, with each question seeking to gain insight into the ability of hybrid BARS methodology to generate valid and reliable data and survey instruments. The first examined the proposition that occupational competencies, generated by occupational experts, were viewed as important by a wider cross-section of people within the profession. Accordingly, frequency analysis criteria, used previously in hybrid BARS studies to assess the importance of performance dimensions, were applied to each specific competency. Results showed that 33 of the 37 competencies identified through the BARS process were classified as 'must-have' or 'should-have.' Only four competencies were considered 'unimportant' (with one of these competencies missing 'should-have' classification by one percentage point). The net result of this analysis is that

hybrid BARS procedures do identify occupation-specific competencies which are seen to be important[1].

The second research question explored the classification of competencies into performance dimensions. Although previous hybrid BARS studies have not determined previously the integrity of dimension structures, occasional traditional BARS studies did make attempts. These analyses met with mixed results (see Chapter 2). Invariably, the studies used confirmatory, as opposed to exploratory (as used in this study), factor analysis to determine the integrity of their scales (e.g., Kafry et al., 1979). This difference notwithstanding, one study which undertook comparable techniques to the present research was that of Stoskopf et al. (1992). Using principle component exploratory factor analysis, they analysed 27 items indicative of effective nursing performance, and derived a stable and conceptually meaningful three factor solution using the following criteria:

- a minimum of six significant loadings per factor;
- the elimination of factors which account for less than 5% of the variance; and,
- 75% or more of the explained variance is accountable by the retained factors.

The factor analysis results for the present study did not meet these, or similarly acceptable, criteria. Although the hybrid BARS procedure required occupational experts to group specific competencies representative of effective refereeing performance, which were seen to be related through practical or theoretical links, a clear and stable empirical factor solution was not found. The *importance* data set did produce six interpretable factors, however, closer examination of the initial statistics revealed that insufficient variance was accounted for by the solution (63.4%), and almost half of the variance (48.3%) was accounted for by two factors. Conversely, the small amounts of variance accounted for by Factors 4 to 6 was indicative of unique variance dominating the common variance structure.

Moreover, the factor solution derived from the *importance* data was only marginally similar to the dimension structure developed from the BARS procedure. Only Factor 6 matched identically a performance dimension (Dimension 5). Some factors did contain sub-groups of items emanating from a single performance dimension, however, items from other dimensions were also present. Thus, there was not enough clear differentiation to support the dimension structure uncategorically.

---

[1] Finding were based on criteria that have not been statistically substantiated previously in research literature. To lend greater credibility to these conclusions, more powerful forms of statistical analysis are required. Accordingly, the present study incorporates results presented in Chapter 6 to verify this finding.

Further validation techniques confirmed an unstable factorial structure. Six random split-half principle component factor analyses were conducted on the original data. These analyses failed to demonstrate a common number of extracted factors or accountable variance. Similarly, scree-test criteria applied to full and split-half data sets continued to provide inconsistent results.

Factor analysis of the *preparedness* and *improvement-priority* data sets also failed to confirm the existence of underlying factors or constructs. Analysis of these data followed the same pattern as was used with the *importance* data. Factor analysis results where again unclear and ambiguous, with a variety of differing factor solutions provided by full and split-half data sub-sets.

A second type of analysis, in the form of Rasch scaling, also failed to demonstrate underlying dimensions in the data. The Rasch technique supplied fit statistics and item consistency indexes indicative of unidimensional data for the *importance*, *preparedness* and *improvement-priority* data.

Given the combined factor analysis and Rasch findings, hybrid BARS did not provide a competency classification that was able to be supported through statistical analysis. However, findings do not indicate necessarily that performance dimensions are a not valuable idea for facilitating discussion. In particular, dimensions provide a mechanism for grouping competencies into different themes, thus enabling the data to be explored and more focused discussion to occur.

The third research question set out to assess the reliability of the research instrument. This was achieved in two ways. Firstly, inspection of item consistency indexes (analogous to Cronbach's *alpha* coefficient) for each perspective showed all indexes to be over 0.90. This verified that items contained within each perspective provided high levels of internal consistency for the research instrument. Secondly, internal consistency, measured across the perspectives of *importance*, *preparedness* and *improvement-priority,* were assessed. This reliability measure was calculated through a Pearson's Product Moment Correlation coefficient ($r = 0.93$), and indicated that respondents were interpreting each perspective as intended by the investigator. Consequently, it can be concluded that hybrid BARS instruments do demonstrate acceptable reliability characteristics.

The net result of all the findings presented in this chapter are generally positive. The BARS procedure described accurately the performance domain of elite soccer referees, and the research instrument predicated on specific refereeing competencies demonstrated acceptable levels of reliability. As a consequence, all results in subsequent chapters are presented with competencies as variables in their own right. Accordingly, the following chapter explores, in greater detail, the hypotheses and research questions from Themes 3 and 4. Specifically,

results focus on the order of competencies within each perspective, thus facilitating examination of possible significant differences both between and within perspectives.

CHAPTER SIX

THE IMPORTANCE AND PREPAREDNESS OF REFEREEING COMPETENCIES

Introduction

The previous chapter utilised survey data to address issues associated with hybrid BARS processes and instruments. Results in this chapter shift the focus of findings away from specific issues related to BARS, and examine competencies as they relate to elite refereeing. In particular, results presented in this chapter have four aims. These are the determination of: competency *importance*; competency *preparedness*; potential differences between the perspectives of *importance* and *preparedness*; and, possible group differences existing within each perspective.

In accordance with these four aims, the presentation of results within this chapter is divided into four sections. The first two sections assess the relative *importance* and *preparedness* of each competency. The remaining two sections answer a number of research hypotheses concerning the possible existence of perspective and group differences. In combination, the findings presented in this chapter address research questions and hypotheses from Theme 3 and Theme 4.

THE RELATIVE IMPORTANCE OF REFEREEING COMPETENCIES

As described in Chapter 3, the identification of essential refereeing competencies, developed via the hybrid BARS technique, made no attempt to delineate the relative importance of each competency. However, it can be argued that all competencies are not of equal importance (despite the *essential* nature of identified competencies). This argument is based on the diversity of competencies developed in the hybrid BARS process, and the relative necessity of each competency to the referee's role. Accordingly, discussion in this section examines results relating to the research question:

How are essential refereeing performance criteria perceived by soccer stakeholders in terms of relative importance?

To address this question, information is drawn from the *importance* data set. Two analytic techniques, namely, percentage frequency analysis and Rasch scaling, are utilised to specify a hierarchy of competencies.

Percentage Frequency Analysis

The stratification of competency *importance* was flagged initially in the previous chapter. Here, the competencies were classified into categories of 'must-have', 'should-have' and 'unimportant.' Consequently, the classification categories outlined in Chapter 5 (see Table 5.6) presented a simplistic hierarchy, which gave a general indication of the relative *importance* of competencies.

Notwithstanding the categorisation of competencies, classifications of this type do not specify a detailed *importance* hierarchy. For example, questions remained concerning the degree of equality among competencies within a single classification category. However, closer scrutiny of the frequency analysis results helps clarify this issue to some degree. This is achieved through using percentage frequency counts as a numerical foundation for determining *importance* ranking. By using the 'should-have' criteria as a benchmark[1] (i.e., percentage of respondents which indicated a response category of 1, 2 or 3), the original three classification categories can be disregarded. A clearer hierarchy of competencies then becomes apparent. The rank order of competencies according to this revised criteria is shown in Table 6.1.

From competencies listed in Table 6.1, those related to the decision-making role of the referee are seen to be the most important (e.g., competencies 2.1, 2.2., 2.4 , and 6.4). These competencies were classified formerly into the 'must-have' category. However, a number of other 'must-have' competencies are not placed highly in the ranking hierarchy. Competency 1.6, *works as a team with assistant referees*, is an obvious example. This competency is ranked 16th using frequency analysis counts. The finding indicates potential differentiation in the relative importance of competencies within the same classification categories.

At the lower end of the *importance* continuum, four competencies which were classified in Chapter 5 as 'unimportant' occupied the four lowest ranking positions using the revised method. However, the ranking positions for these lower ranked competencies were not shared. Again, this finding adds further evidence to a competency hierarchy. As this differentiation is consistent with other competencies within the *importance* continuum, it is evident that the assessment of competency importance becomes more distinct when competencies are removed from the three-category frequency analysis classifications used in Chapter 5.

---

[1] This benchmark was selected as it represents the lowest criteria available before competencies were interpreted as 'unimportant.'

Table 6.1 Competencies by Importance ranking – cumulative percentages

Competency		Percentage	Rank
1.1	understands and interprets the Laws of Soccer correctly	99	1
2.1	applies the Laws of Soccer consistently (within each game and over the season)	99	1
2.2	observes, analyses and correctly interprets incidents	99	1
2.3	reacts quickly and effectively to incidents	99	1
2.4	distinguishes between fair and foul play	99	1
2.9	moves to obtain optimum positions (i.e., place with best view and close enough to react effectively)	99	1
6.1	maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)	99	1
6.4	maintains concentration during the game	99	1
1.2	observes incidents and decides on appropriate action (e.g., fouls, player acting, free kick, booking, play-on)	98	9
1.7	observes play from the best position	98	9
2.5	interprets and discriminates between the severity of fouls	98	9
2.7	distinguishes between advantage and disadvantage	98	9
2.8	manages conflict (communication with players, use of presence and personality)	98	9
4.1	applies the Laws and sanctions (e.g., free kick, yellow and red cards)	98	9
4.4	monitors player behaviour	98	9
1.6	works as a team with assistant referees	97	16
3.4	communicates (verbally and non-verbally) with players on and off the field	97	16
4.2	displays positive behaviours and attitudes	97	16
4.5	manages all aspects of the game (e.g., match control/players)	97	16
1.4	anticipates players actions and reactions (e.g., retaliation to over physical play)	95	20
1.5	keeps a complete record of the game (e.g., bookings, goals, substitutions)	95	20
6.2	maintains appropriate levels of personal health (e.g., correct diet, weight control)	95	20
6.3	undertakes appropriate risk management procedures (e.g., know legal responsibilities, state of the pitch)	95	20
3.2	communicates decisions with clear hand signals	94	24
4.3	manages disputes between players/coaches and match officials	94	24
3.3	undertakes report writing and record keeping (e.g., send off reports, administrative reports)	92	26
3.5	mediates disputes between opposing players	92	26
3.6	communicates confidently with assistant referees	92	26
3.1	effectively uses the whistle (e.g., volume, tone, timing, length, player reaction)	91	29
5.1	prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared)	91	29
2.6	encourages attacking play	90	31
3.7	communicates appropriately with coaches	90	31
5.4	manages personal anxiety	90	31
5.2	undertakes mental preparation for the match (e.g., visualisation)	89	34
3.8	communicates with referee's inspectors (e.g., post match discussions, self reflection, constructive criticism)	86	35
1.3	understands how the game is played (e.g., tactics and strategy, analyse patterns of play)	82	36
5.3	engages in post match activities (e.g., talk to coaches, attend post game functions)	74	37

Notwithstanding the success of percentage frequency counts in delineating an importance continuum, a number of competencies acquired the same ranking position. This is highlighted in Table 6.1, where over half the items (n=24) cluster in the top four ranking positions. Although differences in rankings are more obvious at the lower end of the continuum, it is apparent the separation of competencies is insufficient to delineate an unambiguous importance continuum. Such lack of ranking independence is of some concern, as the degree of success in defining a continuum depends on the extent to which the items are separated (Wright & Masters, 1982). In relation to this concept, competencies must be sufficiently well separated to identify direction and meaning. When the importance of competencies are viewed across the entire *importance* continuum, frequency analysis procedures are insufficient. Moreover, the process does not provide any useful indices, empirical or otherwise, on which substantive judgements and conclusions can be based.

However, the application of Rasch analysis overcomes the statistical and interpretive concerns described above. As discussed in Chapter 4, Rasch analysis produces estimates of item difficulty (analogous to competency *importance* for the present data) on a logit scale. Moreover, the spread of competencies on the scale can be assessed by the 'reliability of estimate' (Adams & Khoo, 1996). Consequently, the suitability of the continuum, in terms of item separation and separation stability, can be determined.

Rasch Analysis

Using the *Quest* software package (Adams & Khoo, 1996), the *importance* data were submitted to Rasch analysis. The fit statistics from this scaling process were discussed initially in Chapter 5 in relation to data homogeneity (see Table 5.8). However, it is important to review the acceptable item fit statistics here. A reported infit mean square of 0.99, and an infit t of -0.22 showed data were suitable for Rasch modelling. A reliability estimate of 0.88 indicated a stable dispersion of competencies across the *importance* continuum.

Table 6.2 provides each competency's item estimate. Additionally, it lists each competency according to its *importance* rank derived from the item estimates. The hierarchy of competencies displayed in this table, coupled with the preceding reliability estimate, allows meaningful discussion and firm conclusions to be made about the relative importance of each competency.

A number of pertinent issues arise from Table 6.2, and are worthy of further examination. Firstly, the competency ranked as most important, i.e., competency 1.1, *understands and interprets the Laws of Soccer correctly*, represents the most important refereeing competency. This conclusion can be drawn as the competency is the only one with a positive

Table 6.2 Competencies by Importance ranking – Rasch analysis

	Competency	Estimate	Rank
1.1	understands and interprets the Laws of Soccer correctly	1.32	1
1.2	observes incidents and decides on appropriate action (e.g., fouls, player acting, free kick, booking, play-on)	0.80	2
2.1	applies the Laws of Soccer consistently (within each game and over the season)	0.76	3
2.2	observes, analyses and correctly interprets incidents	0.6	4
6.4	maintains concentration during the game	0.56	5
4.1	applies the Laws and sanctions (e.g., free kick, yellow and red cards)	0.51	6
2.4	distinguishes between fair and foul play	0.49	7
2.8	manages conflict (communication with players, use of presence and personality)	0.47	8
6.1	maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)	0.45	9
1.6	works as a team with assistant referees	0.43	10
2.5	interprets and discriminates between the severity of fouls	0.36	11
2.7	distinguishes between advantage and disadvantage	0.33	12
1.7	observes play from the best position	0.3	13
2.3	reacts quickly and effectively to incidents	0.26	14
4.4	monitors player behaviour	0.26	14
2.9	moves to obtain optimum positions (i.e., place with best view and close enough to react effectively)	0.25	16
4.5	manages all aspects of the game (e.g., match control/players)	0.17	17
1.5	keeps a complete record of the game (e.g., bookings, goals, substitutions)	0.09	18
3.4	communicates (verbally and non-verbally) with players on and off the field	0.07	19
4.2	displays positive behaviours and attitudes	0.04	20
1.4	anticipates players actions and reactions (e.g., retaliation to over physical play)	-0.10	21
3.6	communicates confidently with assistant referees	-0.21	22
6.2	maintains appropriate levels of personal health (e.g., correct diet, weight control)	-0.30	23
6.3	undertakes appropriate risk management procedures (e.g., know legal responsibilities, state of the pitch)	-0.33	24
2.6	encourages attacking play	-0.35	25
3.3	undertakes report writing and record keeping (e.g., send off reports, administrative reports)	-0.35	25
4.3	manages disputes between players/coaches and match officials	-0.35	25
5.1	prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared)	-0.35	25
3.2	communicates decisions with clear hand signals	-0.46	29
3.5	mediates disputes between opposing players	-0.49	30
5.4	manages personal anxiety	-0.58	31
5.2	undertakes mental preparation for the match (e.g., visualisation)	-0.6	32
3.1	effectively uses the whistle (e.g., volume, tone, timing, length, player reaction)	-0.61	33
3.8	communicates with referee's inspectors (e.g., post match discussions, self reflection, constructive criticism)	-0.61	33
3.7	communicates appropriately with coaches	-0.7	35
1.3	understands how the game is played (e.g., tactics and strategy, analyse patterns of play)	-0.81	36
5.3	engages in post match activities (e.g., talk to coaches, attend post game functions)	-1.34	37
	Mean	0.0	
	Standard Deviation	0.52	

item estimate greater than two standard deviations from the mean.

Secondly, and more generally, all competencies in the top half of the continuum relate to activities that engage the referee during the course of a game. Competency 6.1, *maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)*, would appear to be the only exception to this finding (although it could be argued that proficiency on this competency is fundamental to the successful execution of many other important competencies). Also, the seven top-ranked competencies are associated with the application and interpretation of rules, i.e., decision-making. This result is consistent with the categorisation, and ranking, of competencies using percentage frequency analysis counts.

Thirdly, only one of the eight 'communication' competencies is ranked highly. This item, i.e., competency 2.8, *manages conflict (communication with players, use of presence and personality),* was ranked 8th. The next most important communication item, competency 3.4, *communicates (verbally and non-verbally) with players on and off the field*, was ranked 19th. These results are similar to frequency analysis findings, which categorised numerous communication competencies into the 'should-have,' as opposed to the 'must-have,' category.

At the lower end of the *importance* continuum, competencies of note which ranked lowly include communication with coaches (competency 3.7), communication with referee inspectors (competency 3.8), and an understanding of game tactics and strategies (competency 1.3). However, one competency, i.e., 5.3, *engages in post match activities (e.g., talk to coaches, attend post game functions)*, has a negative item estimate more than two standard deviation from the mean.

As the focus of this section is to determine the relative importance of refereeing competencies, it is appropriate to subject competency 5.3, and the next lowest ranked competency, i.e., 1.3, *understands how the game is played (e.g., tactics and strategy, analyse patterns of play)*, to a paired-sample one-tailed *t* test (*df*=172). This helps to clarify the *importance* status of competency 5.3. The *t* test utilises item estimates (analogous to means) and standard error scores (analogous to standard deviations) obtained from the Rasch analysis for calculations. Based on the formula below, this test establishes if differences between the two competencies is significant[1].

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{sd(\overline{X}_1)^2 + sd(\overline{X}_2)^2}}$$

---

[1] Successive *t* tests between each adjacent pair of skills was rejected due the likelihood of escalating Type 1 errors.

Subsequently, a t value of 3.9 was calculated ($p \geq 0.001$). The significant result obtained from this calculation, when viewed in combination with the standard deviation findings and frequency analysis results, brings into question the relevance of competency 5.3 as an important refereeing competency.

In summary, the findings in this chapter have demonstrated that refereeing competencies identified are not of equal importance. This conclusion was drawn initially through the classification of competencies using a cumulative frequency analysis technique, although the spread of competencies was insufficient for unequivocal conclusions to be drawn. Subsequent item analysis, using Rasch scaling, verified the *importance* hierarchy.

Discussion and Implications

The various forms of analysis used in this section were employed to answer Research Question 3.1: *How are essential refereeing performance criteria perceived by soccer stakeholders in terms of relative importance?* Generally, the results established that competencies are not perceived to be of equal importance. This finding became clearer once more sophisticated approaches to data analysis were applied. In particular, Rasch analysis provided the most comprehensive indication that competencies were spread across an importance continuum.

Specifically, the results provide evidence that competencies which are demonstrable on the field-of-play are seen to be the most important. Of these, competencies which constitute the decision-making aspects of the referee's role received the highest rankings. However, it should be remembered that these results are *relative* rankings, and, as such, do not imply that competencies at the lower end of the continuum are not important. Nonetheless, these findings are unambiguous, although not unexpected given the stated importance of decision-making in the officiating literature.

In contrast to the high rankings associated with decision-making competencies, were the relatively lower rankings given to communication competencies. These latter rankings demonstrated inconsistency with the acknowledged emphasis in the literature on communication competencies. The two lowest ranked communication items, which relate specifically to communication with referee inspectors and coaches, were ranked in equal 33rd and 35th positions, respectively. Two interconnected reasons may be offered for these rankings. Firstly, communication with referee inspectors and coaches does not have an immediate impact on the game. The coach and inspector are removed from the on-field aspects of play, hence, communication with these people takes place, most commonly, at the game's conclusion. Accordingly, the impact of communication on other refereeing competencies, such as consistency, observation, and rule application is negligible. Secondly, the higher proportion of players in the global sample may have accentuated the

result. If, indeed, competencies which immediately effect players are seen to be most important, competencies removed from this process – communication or otherwise – may be perceived to be of lesser importance. This second issue is considered later in this chapter when responses of different sub-groups are discussed.

Notwithstanding these 'relative importance' findings, the only competency that can be dismissed in terms of its relevance to elite soccer refereeing is competency 5.3, *engages in post match activities (e.g., talk to coaches, attend post game functions).* The rejection of this competency is based on three inter-connected statistical reasons. Firstly, Rasch analysis scaled the competency significantly lower than the next lowest ranked competency; secondly, the competency was ranked last by the Rasch analysis and frequency analysis percentage counts; and thirdly, the competency was one of four competencies listed as 'unimportant' by the frequency analysis classification process. While each reason, in its own right, would not be sufficient to dismiss competency 5.3, the combination of these three findings provides substantial evidence for the competency's rejection as an essential competency for elite soccer referees.

The findings presented in this section have implications also for the analysis of BARS data. In particular, frequency analysis (as used in previous hybrid BARS studies, see Anshel, 1995; Anshel et al., 1987; Jessup, 1994; Moore et al., 1997) has proven to be inappropriate for defining items on a continuum. Competencies tended to be clustered, thus giving a false impression of relative importance. Reasons for the clustering may be found in procedures used to account for item importance. Specifically, the frequency analysis technique took into account the middle Likert scale response category. As such, respondents who were not committed strongly to a competency's importance were included as contributing to item importance. While the collapsing of responses is practically workable for the *classification* of items, it does not provide an unambiguous order of competency importance.

However, Rasch analysis proved to be superior to frequency analysis in two ways. Firstly, it discriminated adequately items along a continuum, thus allowing a finer analysis of item position. Secondly, it was able to produce an index which reflected the stability of items on the continuum (0.88). These types of measures are of fundamental significance, particularly if substantive comparisons and conclusions are to be drawn concerning the relative merits of particular competencies.

More generally, the findings added evidence to support the use of expert panels to identify important occupational competencies. As noted previously, the finding presented in this section show clearly that one competency only, i.e., competency 5.3, *engages in post match activities (e.g., talk to coaches, attend post game functions)*, was significantly less important than all other competencies. Given the improved statistical rigour of Rasch analysis over frequency analysis techniques, it can be concluded with greater certainty here that

competencies, identified by expert panels as essential, are also perceived as important by a wider cross-section of the occupational-role under investigation.

THE RELATIVE PREPAREDNESS OF REFEREEING COMPETENCIES

The presentation of results in this section mirrors the format presented for the *importance* data., i.e., percentage frequency analysis and Rasch scaling. In doing so, analysis provides an overall picture of the level of preparedness referees exhibit when performing the competencies under investigation. Subsequent results answer the research question:

How are essential refereeing performance criteria perceived by soccer stakeholders in terms of relative preparedness?

Percentage Frequency Analysis – Preparedness

By following cumulative percentage classification criteria as per *importance* data, all competencies were classified into 'very-well,' 'well,' and 'poorly' prepared categories. Table 6.3 summarises these results (for detailed cumulative percentage counts, see Appendix 17). As shown in the table, a majority of competencies (n=26) were classified into the 'poorly' prepared category, with the remaining competencies classified as 'well' prepared. No competency meet the 'very-well' prepared criteria.

Of those competencies that were classified into the 'well' prepared category, no consistent pattern was identified. The competencies were diverse in their nature (e.g., decision-making, communication, mechanics), and in the context in which the competency was performed (e.g., on-field or off-field). Similarly, competencies classified into the 'poorly' prepared category were not able to be grouped either. They were drawn from across the spectrum of refereeing responsibilities and roles.

Notwithstanding the absence of a pattern in the findings, the classification of competencies across two categories did indicate a possibility that competencies are not undertaken with equal preparedness. However, conclusive evidence is not apparent, as all competencies fell into two categories. Thus, clear discrimination among competencies, particularly with respect to those competencies listed in the 'poorly' prepared category, were not evident. In this regard, results obtained from the categorisation of *preparedness* data is reflective of results obtained for *importance* data set.

Table 6.3 Frequency analysis classification of items with respect to Preparedness

	Competency	Classification		
		Very-well	Well	Poorly
1.1	understands and interprets the Laws of Soccer correctly			✓
1.2	observes incidents and decides on appropriate action (e.g., fouls, player acting, free kick, booking, play-on)			✓
1.3	understands how the game is played (e.g., tactics and strategy , analyse patterns of play)			✓
1.4	anticipates players actions and reactions (e.g., retaliation to over physical play)			✓
1.5	keeps a complete record of the game (e.g., bookings, goals, substitutions)		✓	
1.6	works as a team with assistant referees			✓
1.7	observes play from the best position		✓	
2.1	applies the Laws of Soccer consistently (within each game and over the season)			✓
2.2	observes, analyses and correctly interprets incidents			✓
2.3	reacts quickly and effectively to incidents		✓	
2.4	distinguishes between fair and foul play			✓
2.5	interprets and discriminates between the severity of fouls			✓
2.6	encourages attacking play			✓
2.7	distinguishes between advantage and disadvantage			✓
2.8	manages conflict (communication with players, use of presence and personality)			✓
2.9	moves to obtain optimum positions (i.e., place with best view and close enough to react effectively)		✓	
3.1	effectively uses the whistle (e.g., volume, tone, timing, length, player reaction)		✓	
3.2	communicates decisions with clear hand signals		✓	
3.3	undertakes report writing and record keeping (e.g., send off reports, administrative reports)		✓	
3.4	communicates (verbally and non-verbally) with players on and off the field			✓
3.5	mediates disputes between opposing players			✓
3.6	communicates confidently with assistant referees		✓	
3.7	communicates appropriately with coaches			✓
3.8	communicates with referee's inspectors (e.g., post match discussions, self reflection, constructive criticism)			✓
4.1	applies the Laws and sanctions (e.g., free kick, yellow and red cards)			✓
4.2	displays positive behaviours and attitudes			✓
4.3	manages disputes between players/coaches and match officials			✓
4.4	monitors player behaviour			✓
4.5	manages all aspects of the game (e.g., match control/players)			✓
5.1	prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared)		✓	
5.2	undertakes mental preparation for the match (e.g., visualisation)			✓
5.3	engages in post match activities (e.g., talk to coaches, attend post game functions)			✓
5.4	manages personal anxiety			✓
6.1	maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)		✓	
6.2	maintains appropriate levels of personal health (e.g., correct diet, weight control)			✓
6.3	undertakes appropriate risk management procedures (e.g., know legal responsibilities, state of the pitch)		✓	
6.4	maintains concentration during the game			✓

Closer inspection of the frequency analysis percentage counts (according to the cut-off criteria for the 'well' prepared classification) delineated a clearer hierarchy of competencies (see Table 6.4). Of the 37 competencies examined, 15 competencies received individual ranking positions.

The competencies with the highest rankings were competencies 2.9, *moves to obtain optimum positions (i.e., place with best view and close enough to react effectively)*, and 6.1, *maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)*. A practical relationship exists between these two competencies. Primarily, for a referee to keep pace with play and be in the optimum position to make decisions, a high level of fitness is required. Therefore, given the high ranking of competency 2.9, it is not unexpected that competency 6.1 is ranked similarly.

Of the remaining 11 competencies ranked 10 or higher, five can be viewed as 'mechanical' aspects of refereeing, i.e., the competency's execution is standardised and/or predetermined, and is unrelated to the more complex decision-making, interpretive, and communication requirements of referees. No mechanical competency was ranked lower than 8th position. For example, competency 3.1, *effectively uses the whistle (e.g., volume, tone, timing, length, player reaction)*, was ranked third. Other mechanical-related competencies that were ranked similarly are (in ascending order) 5.1, 1.5, 6.3, and 3.3.

At the opposite end of the *preparedness* continuum, three of the five competencies ranked lowest (i.e. 2.8, 3.4, and 4.4) relate to various forms of interaction between referees and players. This result indicates that the forms of interaction implicit in these competencies are aspects of the referee's role that require additional preparation.

However, despite the emerging *preparedness* hierarchy in Table 6.4, 22 competencies are ranked with one or more other competencies. This occurrence is apparent across seven ranking positions. Given this circumstance, the clarity of the competency hierarchy for the *preparedness* data is insufficient for substantive conclusions to be made. Consequently, Rasch analysis is applied again to determine if an unambiguous *preparedness* hierarchy can be identified.

Table 6.4 Competencies by Preparedness ranking – cumulative percentages

	Competency	Prep %	Rank
2.9	moves to obtain optimum positions (i.e., place with best view and close enough to react effectively)	94	1
6.1	maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)	94	1
3.1	effectively uses the whistle (e.g., volume, tone, timing, length, player reaction)	94	1
5.1	prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared)	93	4
3.2	communicates decisions with clear hand signals	92	5
1.5	keeps a complete record of the game (e.g., bookings, goals, substitutions)	92	5
3.6	communicates confidently with assistant referees	91	7
1.7	observes play from the best position	90	8
2.3	reacts quickly and effectively to incidents	90	8
6.3	undertakes appropriate risk management procedures (e.g., know legal responsibilities, state of the pitch)	90	8
3.3	undertakes report writing and record keeping to maintain a complete record of the game (e.g., send off reports, administrative reports)	90	8
6.2	maintains appropriate levels of personal health (e.g., correct diet, weight control)	88	12
6.4	maintains concentration during the game	87	13
1.1	understands and interprets the Laws of Soccer correctly	86	14
5.2	undertakes mental preparation for the match (e.g., visualisation)	86	14
1.2	observes incidents and decides on appropriate action (e.g., fouls, player acting, free kick, booking, play-on)	85	16
2.4	distinguishes between fair and foul play	85	16
5.4	manages personal anxiety	85	16
4.2	displays positive behaviours and attitudes	84	19
4.1	applies the Laws and sanctions (e.g., free kick, yellow and red cards)	83	20
4.5	manages all aspects of the game (e.g., match control/players)	82	21
1.6	works as a team with assistant referees	80	22
2.7	distinguishes between advantage and disadvantage	80	22
3.5	mediates disputes between opposing players	80	22
3.8	communicates with referee's inspectors (e.g., post match discussions, self reflection, constructive criticism)	80	22
2.2	observes, analyses and correctly interprets incidents	79	26
2.6	encourages attacking play	79	26
1.4	anticipates players actions and reactions (e.g., retaliation to over physical play)	78	28
2.5	interprets and discriminates between the severity of fouls	77	29
4.3	manages disputes between players/coaches and match officials	76	30
3.7	communicates appropriately with coaches	75	31
2.1	applies the Laws of Soccer consistently (within each game and over the season)	73	32
3.4	communicates (verbally and non-verbally) with players on and off the field	73	32
4.4	monitors player behaviour	73	32
5.3	engages in post match activities (e.g., talk to coaches, attend post game functions)	73	32
2.8	manages conflict (communication with players, use of presence and personality)	72	36
1.3	understands how the game is played (e.g., tactics and strategy , analyse patterns of play)	61	37

Rasch Analysis

Submission of all 37 competencies from the *preparedness* data set to Rasch analysis yielded both acceptable fit statistics (see Table 5.11) and an acceptable item reliability estimate (0.88). Item thresholds obtained through this analysis, coupled with associated rankings, are detailed in Table 6.5.

From Table 6.5, the competency perceived to be undertaken with the highest levels of preparation is competency 1.5, *keeps a complete record of the game (e.g., bookings, goals, substitutions)*. This competency is the only item to record an item estimate (1.14) that is more than two standard deviations from the mean at the high-ranking end of the *preparedness* continuum.

More generally, the seven highest ranked competencies listed in Table 6.5 are indicative of competencies which are not associated with rule application and interpretation (although three competencies – ranked in positions 5 to 7 – are a consequence of a rule administration, yet not implicated directly in the rule administration process). For example, competency 5.1, *prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared)*, and competency 6.1, *maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)*, are undertaken away from the competitive game context. Two other competencies, 1.5, *keeps a complete record of the game (e.g., bookings, goals, substitutions)*, and 3.3, *undertakes report writing and record keeping (e.g., send off reports, administrative reports)*, are administrative requirements essentially undertaken during and after the game. It is not until competency 1.1, *understands and interprets the Laws of Soccer correctly*, ranked in 8th position, that a competency connected to the interpretive aspects of the referee's role becomes apparent.

For competencies at the lower end of the ranking continuum, no consistent pattern was identified. A range of competencies are evident, and these relate to player management, communication, rule application, and strategic understanding of the game. Of these, the lowest ranking was given to competency 3.1, *understands how the game is played (e.g., tactics and strategy, analyse patterns of play.* However, as this competency is not ranked highly in terms of *importance*, the low rankings for *preparedness* is possibly of minor practical concern.

In summary, the Rasch scaling and frequency analysis show a continuum of *preparedness* is evident. Percentage frequency analysis provided initial evidence that this was so, with subsequent Rasch analysis verifying the hierarchy.

Table 6.5 Competencies by Preparedness ranking – Rasch analysis

Competency		Estimate	Rank
1.5	keeps a complete record of the game (e.g., bookings, goals, substitutions)	1.14	1
5.1	prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared)	0.82	2
6.1	maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)	0.57	3
3.3	undertakes report writing and record keeping (e.g., send off reports, administrative reports)	0.56	4
3.2	communicates decisions with clear hand signals	0.49	5
3.6	communicates confidently with assistant referees	0.43	6
3.1	effectively uses the whistle (e.g., volume, tone, timing, length, player reaction)	0.42	7
1.1	understands and interprets the Laws of Soccer correctly	0.32	8
6.3	undertakes appropriate risk management procedures (e.g., know legal responsibilities, state of the pitch)	0.31	9
6.4	maintains concentration during the game	0.25	10
2.9	moves to obtain optimum positions (i.e., place with best view and close enough to react effectively)	0.19	11
6.2	maintains appropriate levels of personal health (e.g., correct diet, weight control)	0.13	12
1.7	observes play from the best position	0.12	13
2.3	reacts quickly and effectively to incidents	0.12	13
4.2	displays positive behaviours and attitudes	0.1	15
4.1	applies the Laws and sanctions (e.g., free kick, yellow and red cards)	0.03	16
2.4	distinguishes between fair and foul play	0.02	17
5.2	undertakes mental preparation for the match (e.g., visualisation)	0.01	18
4.5	manages all aspects of the game (e.g., match control/players)	-0.04	19
1.2	observes incidents and decides on appropriate action (e.g., fouls, player acting, free kick, booking, play-on)	-0.1	20
1.6	works as a team with assistant referees	-0.15	21
2.7	distinguishes between advantage and disadvantage	-0.16	22
2.2	observes, analyses and correctly interprets incidents	-0.21	23
2.5	interprets and discriminates between the severity of fouls	-0.21	23
3.8	communicates with referee's inspectors (e.g., post match discussions, self reflection, constructive criticism)	-0.26	25
2.6	encourages attacking play	-0.28	26
4.3	manages disputes between players/coaches and match officials	-0.28	26
5.4	manages personal anxiety	-0.28	26
1.4	anticipates players actions and reactions (e.g., retaliation to over physical play)	-0.29	29
4.4	monitors player behaviour	-0.31	30
3.5	mediates disputes between opposing players	-0.35	31
3.7	communicates appropriately with coaches	-0.38	32
3.4	communicates (verbally and non-verbally) with players on and off the field	-0.43	33
2.1	applies the Laws of Soccer consistently (within each game and over the season)	-0.45	34
2.8	manages conflict (communication with players, use of presence and personality)	-0.45	34
5.3	engages in post match activities (e.g., talk to coaches, attend post game functions)	-0.51	36
1.3	understands how the game is played (e.g., tactics and strategy, analyse patterns of play)	-0.91	37
	Mean	0.0	
	Standard Deviation	0.413	

Discussion and Implications

Results presented in this section addressed Research Question 3.2: *How are essential refereeing performance criteria perceived by soccer stakeholders in terms of relative preparedness?* In general, results demonstrated that refereeing competencies are not perceived as being undertaken with equal preparedness.

Specifically, results derived from frequency analysis and Rasch scaling are pertinent. In particular, frequency analysis techniques indicated the level of competency preparation was not of a capable standard. The high number of competencies listed in the 'poorly' *prepared* category, and the absence of competencies classified into the 'very-well' *prepared* category, support this view. Moreover, Rasch analysis showed highly ranked competencies for *preparedness* were, in fact, peripheral to the central decision-making and arbitration roles of the referee. In combination, these two points raise the possibility that referees are not being prepared adequately for undertaking fundamental officiating competencies.

This latter point is reinforced by competencies at the lower end of the *preparedness* continuum. In particular, the low rankings given to four competencies, i.e., 2.1, *applies the Laws of Soccer consistently* (ranking = 34); 2.7, *distinguishes between advantage and disadvantage* (ranking = 22); 2.2, *observes, analyses and correctly interprets incidents* (ranking = 23); and, 2.5, *interprets and discriminates between the severity of fouls* (ranking = 23) is of concern given the stated importance in the literature of the need for consistent decision-making. A direct comparison between perspectives, using rankings derived from the Rasch analysis, may shed more light on prospective differences. However, such a comparison is inappropriate at this point, given the rankings were derived from separate Rasch scales (the following section addresses this issue in greater detail).

Nevertheless, evidence to support the dichotomy between *importance* and *preparedness* was also apparent from frequency analysis results. The results presented in Table 6.3 illustrated a contrast to the *importance* data findings (see Table 5.6) with respect to the *type* of classifications received. As results have presented trends which are worthy of further investigation, the differences flagged in this section are pursued through more rigorous forms of parametric analysis, the results of which are discussed in the following section.

PERSPECTIVE DIFFERENCES

The previous sections identified the possibility of major differences existing between the *importance*, and *preparedness*, of competencies. This was evident initially from frequency analysis categorisations. Particularly, the numerous competencies classified into the 'poorly' prepared category for *preparedness*, while being classified also as 'must-have' with respect

to *importance*, crystallised this observation. However, given the ordinal nature of Likert scale data, firm conclusions based on frequency analysis classifications are not appropriate. Moreover, results of significance can not be quantified from this form of data.

Notwithstanding the analytical limitations, potential empirical differences between perspectives can be obtained. Statistical procedures incorporating Rasch scaling (which 'convert' ordinal data to interval scores), *t* tests, and Cook's distance, facilitate this analysis. These techniques, and the order in which they should be applied, were detailed previously in Chapter 4. This serves two specific purposes. The first is to assess the possibility of a significant difference between the perspectives of *importance* and *preparedness*, and thus address the following hypothesis:

Hypothesis 3:

There will be no significant difference between the relative importance of elite soccer refereeing criteria and the preparedness of referees for undertaking these criteria.

The second purpose of the data plan is a function of the potential rejection of this hypothesis. If a significant difference is found, a variety of techniques can be applied to the data to ascertain which competencies are most likely to cause this difference.

Rasch analysis

All *importance* and *preparedness* data were submitted to Rasch analysis on a single scale. This was required to avoid discrepancies in item calibration which would ordinarily occur if item thresholds were obtained from two separate scales. Additionally, the technique provides interval measures, i.e., item estimates, that can be used for subsequent parametric analysis. Conceptually, if a significant difference is to be obtained, the *importance* item estimates and the *preparedness* item estimates will be distinctively located on the same Rasch scale. Figure 6.1 illustrates this potential relationship.

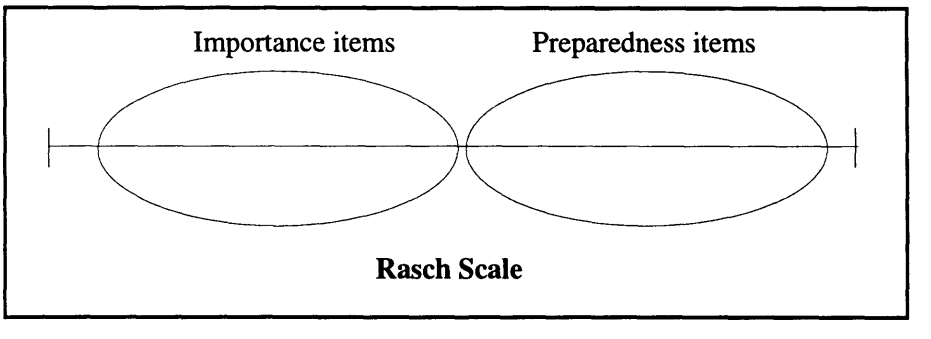

Figure 6.1 Conceptually distinct locations for perspective item estimates

Rasch analysis of all 74 items from the *importance* and *preparedness* data sets yielded acceptable fit statistics (see Table 6.6). Importantly, the item *reliability of estimate* was 0.96.

This indicates items are sufficiently separated and stable, thus facilitating further parametric analysis.

Table 6.6 Fit-statistics for Importance/Preparedness items

Criteria	Statistic
Item estimates	
infit mean square	0.98
infit *t*	-0.63
reliability of estimate	0.96
Case estimates	
infit mean square	0.93
infit *t*	-0.82
reliability of estimate	0.95

Using perspective means (derived from item estimates), tests for assessing significant differences can be implemented. Specifically, perspective differences can determined via a *t* test, as two sets of data are to be analysed (Crowl, 1993).

t tests

The two-tailed form of a non-independent paired sample *t* test was used (with *importance* and *preparedness* as dependant variables) to determine if means were significantly different. This form of the *t* test was selected as analysis was non-directional, and data were derived from the same group. To facilitate the *t* test, means were generated by re-sorting item estimates into perspectives (see Appendix 18). Means were calculated subsequently and compared for each perspective (*importance*=0.5727, *sd*=0.567; *preparedness*=(-)0.5718, *sd*=0.278). The means indicated an initial difference in perspectives.

A *t*-value of 11.6 (*df*=36) was calculated using the SPSS statistical package. This indicated a significant difference ($p<.001$) between perspectives, and confirmed differences that arose previously from frequency analysis classifications. However, despite a significant difference being found, the *t* test was incapable of determining explicitly which competency, or competencies, had a significant impact in causing this difference. Such specific detail can be obtained using Cook's distance. This technique was described briefly in Chapter 4, and is elaborated upon in the following sub-section.

Cook's Distance

Cook's distance is a mathematical maximisation procedure that is sensitive to data which are different from the rest, i.e., 'outliers' (Stevens, 1996). It measures the influence that one observation exerts on the regression line by summarising the changes in residuals that result from deleting that observation from the regression set (Stevens, 1996). In doing so, Cook's

distance summarises the effects of changes in both the vertical and horizontal distances of the residuals from the regression line. Cook's distance has an F-distribution with N-p-1 degrees of freedom (N = the number of cases, and p is the number of parameters in the regression equation). Algebraically, Cook's distance is given by:

$$CD_i = \frac{\left(\hat{\beta} - \hat{\beta}_{(-i)}\right)' X' X \left(\hat{\beta} - \hat{\beta}_{(-i)}\right)}{(k+1) MS_{res}}$$

where $\hat{\beta}_{(-i)}$ is the vector of estimated regression coefficients with the ith data point deleted, k is the number of predictors, and MS_{res} is the residual (error) variance for the full data set.

Given the analytical applications of Cook's distance, the technique was utilised to establish which competencies were implicated significantly in causing the significant perspective difference derived from the t test.

Using item estimates derived for the t test, Cook's distance analysis did not identify specific competencies which were implicated significantly in causing the difference between perspectives (values greater than 4.12 are significant at an alpha level of 0.05 – see Appendix 19 for full details). The relationships among items, along with the relevant regression equation derived from the Cook's distance analysis, is shown in Figure 6.2. From a statistical perspective, the non-significant result suggests all competencies are making an accumulative effect to the significant difference between perspectives.

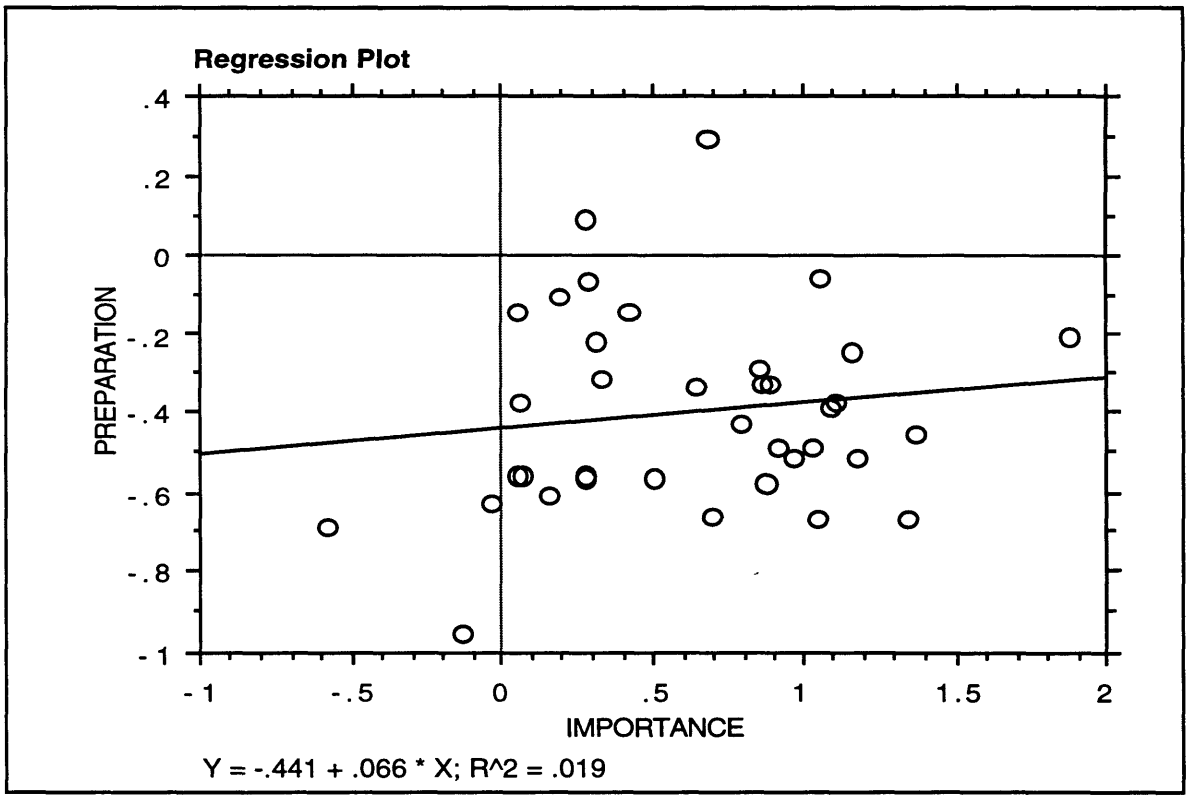

Figure 6.2 Perspective differences regression plot

However, in an attempt to unpack the data, rank-order comparisons between each adjacent pair of competencies, across each perspective, may provide an indication of which competencies are driving the significant difference, e.g., compare the *importance* rank of competency 1.1 against the *preparedness* rank of the same competency. Although rankings are without the statistical and empirical power of parametric techniques (such as Cook's distance), they can provide detail not apparent from empirical procedures.

Rank-Order Comparisons

For determining rank-order comparisons, competencies need to be ranked within each perspective. This is achieved through three specific steps. Firstly, a numerical value for each competency is determined. Item estimates derived from Rasch analysis fulfil this requirement (see Table 6.6 for fit statistics, and Appendix 18 for item estimates). Secondly, using the numerical value established in Step 1, competencies are ranked according to their level of *importance* and *preparedness*, from highest to lowest. Thirdly, a corresponding ranking position is allocated for each competency, and comparisons made across perspectives. Ranking comparisons, for each competency-pair, are then determined by calculating the difference between respective ranking positions.

When these steps are applied to the *importance* and *preparedness* data, discrepancies between perspectives emerge. As can be seen from Figure 6.3, 14 competencies recorded a ranking differential of 15 or more places (a complete ranking inventory is provided in Appendix 20). The largest difference occurred with competency 2.1, *applies the Laws of Soccer consistently (within each game and over the season)*, with a differential of 31 places. Other competencies which are similarly opposed include:

3.1 effectively uses the whistle (e.g., volume, tone, timing, length, player reaction) *(ranking differential = 27)*

2.8 manages conflict (communication with players, use of presence and personality) *(ranking differential = 26)*

3.2 communicates decisions with clear hand signals *(ranking differential = 24)*

5.1 prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared) *(ranking differential = 24)*

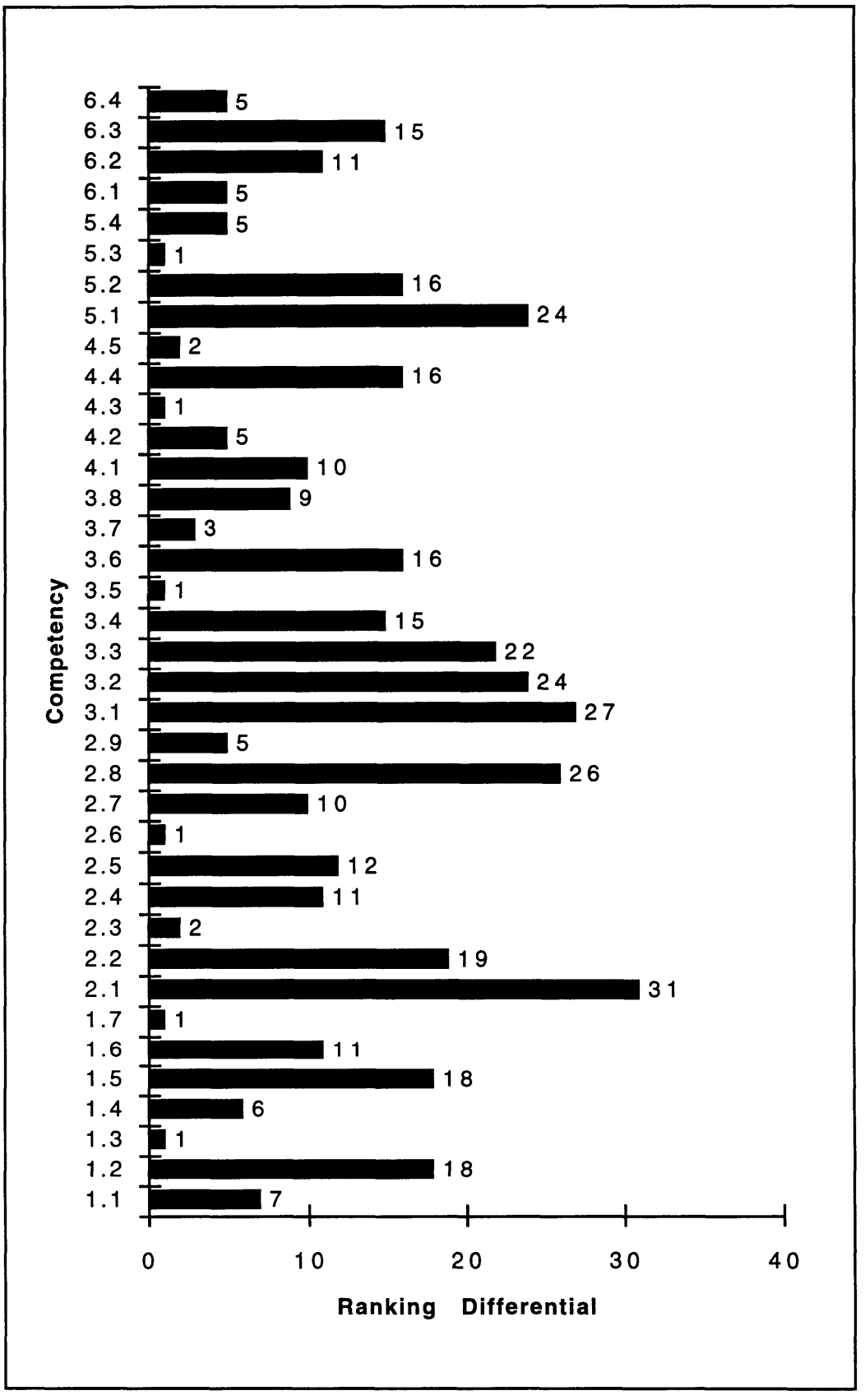

Figure 6.3 Competency rank-order differentials for Importance and Preparedness perspectives

3.3 undertakes report writing and record keeping (e.g., send off reports, administrative reports)
(ranking differential = 22)

All competencies listed above have a ranking differential of 20 or more positions, thus they can be considered, along with competency 2.1, to be involved notably in causing the significant difference between perspectives.

However, from an applied viewpoint, ranking differentials do not reveal the full implications of perspective differences. To illustrate this point, Table 6.7 lists all competencies with ranking differentials of 15 or more places. Additionally, the table highlights those competencies which are ranked highly in *importance*, but ranked low in terms of *preparedness*.

Table 6.7 Marked ranking differentials across perspectives

Competency	Importance Ranking	Preparedness Ranking	Ranking Differential
2.1 applies the Laws of Soccer consistently (within each game and over the season)	3	34	31
3.1 effectively uses the whistle (e.g., volume, tone, timing, length, player reaction)	33	6	27
2.8 manages conflict (communication with players, use of presence and personality)	9	35	26
3.2 communicates decisions with clear hand signals	29	5	24
5.1 prepares well in advance of the match (time of arrival at ground - on time, presentation on arrival, kit prepared)	26	2	24
3.3 undertakes report writing and record keeping (e.g., send off reports, administrative reports)	25	3	22
2.2 observes, analyses and correctly interprets incidents	4	23	19
1.2 observes incidents and decides on appropriate action (e.g., fouls, player acting, free kick, booking, play-on)	2	20	18
1.5 keeps a complete record of the game (e.g., bookings, goals, substitutions)	19	1	18
3.6 communicates confidently with assistant referees	22	6	16
4.4 monitors player behaviour	14	30	16
5.2 undertakes mental preparation for the match (e.g., visualisation)	32	16	16
6.3 undertakes appropriate risk management procedures (e.g., know legal responsibilities, state of the pitch)	24	9	15
3.4 communicates (verbally and non-verbally) with players on and off the field	18	33	15

In practice, those competencies that are not highlighted are of lesser concern (due to their diminished importance), so major ranking differentials, with respect to *preparedness*, are probably of minor consequence. Of greater practical significance are competencies ranked highly for *importance* and lowly for *preparedness* (those shaded in Table 6.7). Examples of these competencies include:

1.2 observes incidents and decides on appropriate action (e.g., fouls, player acting, free kick, booking, play-on);

2.1 applies the Laws of Soccer consistently (within each game and over the season);

2.2 observes, analyses and correctly interprets incidents; and,

2.8 manages conflict (communication with players, use of presence and personality).

When findings are viewed from this 'high importance – low preparation' dichotomy, the results presented in Table 6.7 are consistent with competency frequency analysis classifications presented in the previous sections. Moreover, results such as these begin to reveal, in a constrained manner, perceptions of the standard of refereeing performance at the most elite level of Australian soccer.

Discussion and Implications

Results presented in the preceding discussion confirmed a significant difference ($p<.001$) existed between the *importance* and *preparedness* of competencies. Accordingly, the following null hypothesis is rejected, and the alternative accepted:

Hypothesis 3: *There will be no significant difference between the relative importance of elite soccer refereeing criteria and the preparedness of referees for undertaking these criteria.*

Empirical verification to determine which competencies had a statistically significant effect in causing this difference was not apparent. This indicated *all* competencies had an accumulative effect in the rejection of Hypothesis 3. However, a rank order of competencies provided a useful, although less precise, indication of which competencies most contributed to the difference.

A range of refereeing roles, responsibilities, and skills were implicated by the ranking process. Most particularly, these included competencies associated with rule application, communication, game preparation, and risk management. However, not all these competencies are necessarily seen to be important.

From the practical perspective, this distinction is meaningful, i.e., the more important a competency, the more consequential it becomes to ensure the competency is undertaken with a high degree of preparedness. In this regard, implications arising from four competencies, i.e., 1.2, 2.1, 2.2, and 2.8, should not be understated. In general, all four competencies are related to on-field actions. More specifically, competencies 1.2 and 2.2 are linked to the observation, integration *and* designation of appropriate refereeing decisions and actions, i.e., free-kicks, play-on, etc. Competency 2.1, *applies the Laws of Soccer consistently (within*

each game and over the season) demands these related competencies do not vary over the duration of a game, or from one game to another. The final competency in this group, 2.8, *manages conflict (communication with players, use of presence and personality)*, is interesting given the competency is removed from the decision-making process (in terms of rule interpretation), yet is seen to be fundamental to the smooth running of the game. The physical nature of soccer ensures that conflict, both physical and verbal, occurs from time-to-time. If not managed effectively, conflict can over-ride the primary focus of a match. However, the dichotomy that exists between the *importance* and *preparation* of these four competencies is of marked concern, particularly given the proficient execution of these competencies are, without doubt, central to effective refereeing (see Chapter 1).

A secondary consequence of these finding concerns their implications for the future direction of referee training and development. Results presented in this section have highlighted refereeing competencies which show a contradiction between competency *importance* and *preparedness*. Obviously, competencies which are high in *importance*, and low in *preparedness*, are competencies which should receive most emphasis in referee development, especially at the elite level.

GROUP DIFFERENCES

The perspective differences described in the previous section identified discrepancies between the importance of a competency, and how well referees were prepared in undertaking that competency. Although a significant differences was found, it is not clear if this difference is consistent across groups, i.e., do groups perceive the same competencies as important, and the level of competency preparedness the same? Although not reported previously, the possibility of group differences arose initially in frequency analysis findings (see Appendices 12 and 17). As the present study sought opinion across a sample which, from a soccer viewpoint, is diverse, it is not unreasonable to assume differences of opinion between the officiating and competitive groups within each perspective. Accordingly, the hypotheses that are posed in this section, expressed in their null form, are:

Hypothesis 4.1

There will be no significant difference of opinion about the relative importance of elite refereeing performance criteria among referees assistant referees, referee inspectors, players, and coaches.

Hypothesis 4.2

There will be no significant difference of opinion about the relative preparation of elite refereeing performance criteria among referees, assistant referees, referee inspectors, players, and coaches.

As was the case in determining perspective differences, a range of statistical procedures are used to address these hypotheses (see Figure 4.2). By way of summarised review, these include Rasch analysis, MANOVA and ANOVA, Cook's distance, and rank-order comparisons.

Rasch Analysis

As detailed in Chapter 4, Rasch analysis provides an index of a person's 'ability' by transforming raw scores, from a series of items, and giving that person a score on a single logit scale, i.e., a case estimate. With respect to this study, an individual's placement on the scale is indicative of perceptions of *importance* or *preparedness* for the 37 competencies. Accordingly, case estimates were required for each perspective, thus necessitating the generation of two scales.

Data from each perspective were submitted for scaling. Fit-statistics for each scale were presented previously in Chapter 5 (see Tables 5.13 and 5.18), however, they are summarised in Table 6.8 for convenience.

Table 6.8 Case estimate fit-statistics for Importance and Preparedness data

	Importance Data	Preparedness Data
infit mean square	1.07	1.03
infit *t*	0.00	-0.34
reliability of estimate	0.89	0.96

The statistics illustrate a good fit of data to model. Moreover, the high reliability of estimate measures show that cases (respondents) were sufficiently separated, and their relative positions on the respective scales were stable. Specific case estimates are provided in Appendix 21. Given these acceptable fit-statistics, possible group differences can be explored within each perspective using the derived case estimates.

MANOVA and ANOVA

Inspection of group means, based on case estimates within each perspective, provides initial indication that group differences may be apparent (see Table 6.9). However, to make inferences concerning group differences based on means alone is spurious and can lead to imprecise conclusions. To provide more accurate assessment of group differences, perspective scores were analysed and found to be significantly correlated (Pearson's r=0.224, p=0.001). Subsequently, group means in Table 6.9 were examined further for the purpose of hypothesis testing.

Table 6.9 Case estimate means by Group and Perspective

Group	Means	
	Importance	Preparedness
Officiating group	-1.9854386	-1.4573333
Competitive group	-1.3988991	-0.1611818

Initially, MANOVA techniques were applied to the data utilising case estimates sorted according to group. This form of analysis is imperative in the first instance, primarily to avoid the possibility of escalating Type 1 errors occurring when two dependent variables are analysed. For the MANOVA in the present study, each group represented independent variables, and each perspective represented dependent variables. If significant differences are found using MANOVA, subsequent univariate one-way ANOVA specifies which dependent variable/s are causing the significant difference.

The MANOVA assumptions of multivariate normality, homogeneity of variance/covariance matrices and linearity were tested as part of the analysis. Multivariate normality was tested by inspection of normal probability plots for each dependent variable and plots of pooled residuals after estimating the model, and there was no evidence that the multivariate assumption was violated.

The homogeneity of variance/covariance matrices was tested using the Box M procedure. This yielded a Box M of 37.956, which is significant ($F(3,372066)=37.956$, $p<.0001$). Although the Box M test is sensitive (Tabachnick & Fidell, 1996), the result is of concern because the cell frequencies (58 and 109) are discrepant. Inspection of the variance/covariance matrices revealed that the variance of *importance* for the officiating group was less than the variance of that variable for the competitive group, but that the reverse was the case for the *preparedness* variable. Similarly, inspection of the covariance matrices revealed that the covariances, with the exception of one cell, were greater for the officiating group. As a consequence, the pooled matrix may over- or under-estimate error variance. For this reason conservative alpha levels were used in the MANOVA. Linearity was tested by inspection of scatter plots for the pairs of dependent variables within each cell and these were found to be satisfactory. A summary of subsequent MANOVA results is provided in Table 6.10.

Table 6.10 MANOVA results of significance

Test Name	Value	Exact F	Hypoth. DF	Error DF	Sig. of F
Pillais	.30792	36.48354	2.00	164.00	.000
Hotellings	.44492	36.48354	2.00	164.00	.000
Wilks	.69208	36.48354	2.00	164.00	.000
Roys	.30792				

The results of this analysis show there is a significant difference between groups on one or more dependent variables. Subsequent univariate F-tests (1,165 *df*), detailed in Table 6.11, confirms a significant differences exists between groups for both dependent variables.

Table 6.11 Univariate analysis tests for significance

Variable	Mean Squares	F-value	Significance
Importance	12.7529	12.93501	0.004
Preparedness	48.5551	55.74073	0.001

However, as was the case when significant perspective differences were obtained using *t* tests, the MANOVA and ANOVA procedures were not able to specify which competencies are causing significant group differences. As such, Cook's distance was again employed to ascertain if one or more competencies are significantly implicated in causing group differences.

Cook's Distance

As described previously in this section, Cook's distance is a regression technique that is recommended for identifying outliers that are influential in determining significant differences. For Cook's distance to proceed in the context of this analysis, the production of additional item estimates are required. This enables perceptions of *importance* and *preparedness* to be determined for each group.

Estimates are required from four sub-scales generated from the following data: *importance*: data (2 scales: 1 x competitive group; 1 x officiating group); and, *preparedness* data (2 scales: 1 x competitive group; 1 x officiating group). Relevant fit statistics and item estimates for all four scales from the Rasch analysis are provided in Appendix 22. Item estimates for each group, within each perspective, were compared to determine Cook's distance.

Subsequent analysis using Cook's distance did not show any one competency making statistically significant contributions to group differences for either perspective (values greater than 4.12 are significant at an alpha level of 0.05 – see Appendix 23 for full details). Graphical representation of these findings, with relevant regression equations, is shown in Figures 6.4 and 6.5. Given the results of this analysis, it is concluded that all competencies are making an accumulative contribution to the significant group difference within each perspective.

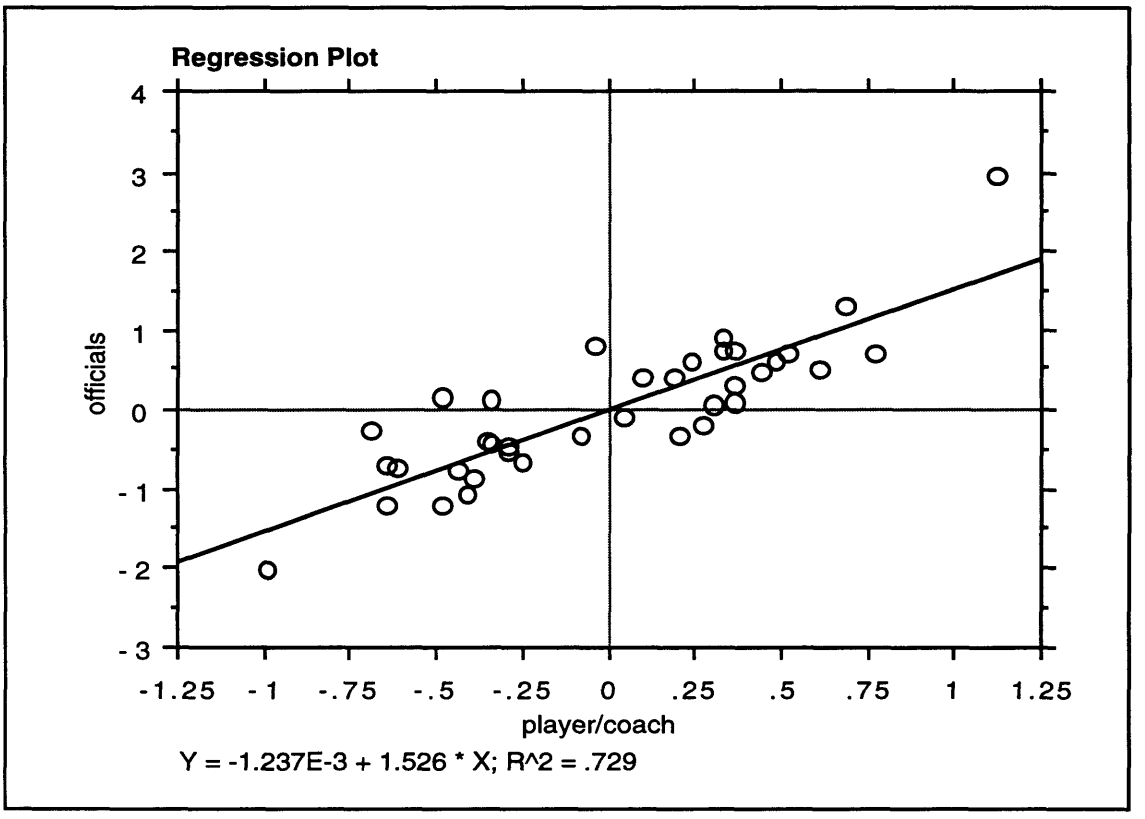

Figure 6.4 Importance regression plot

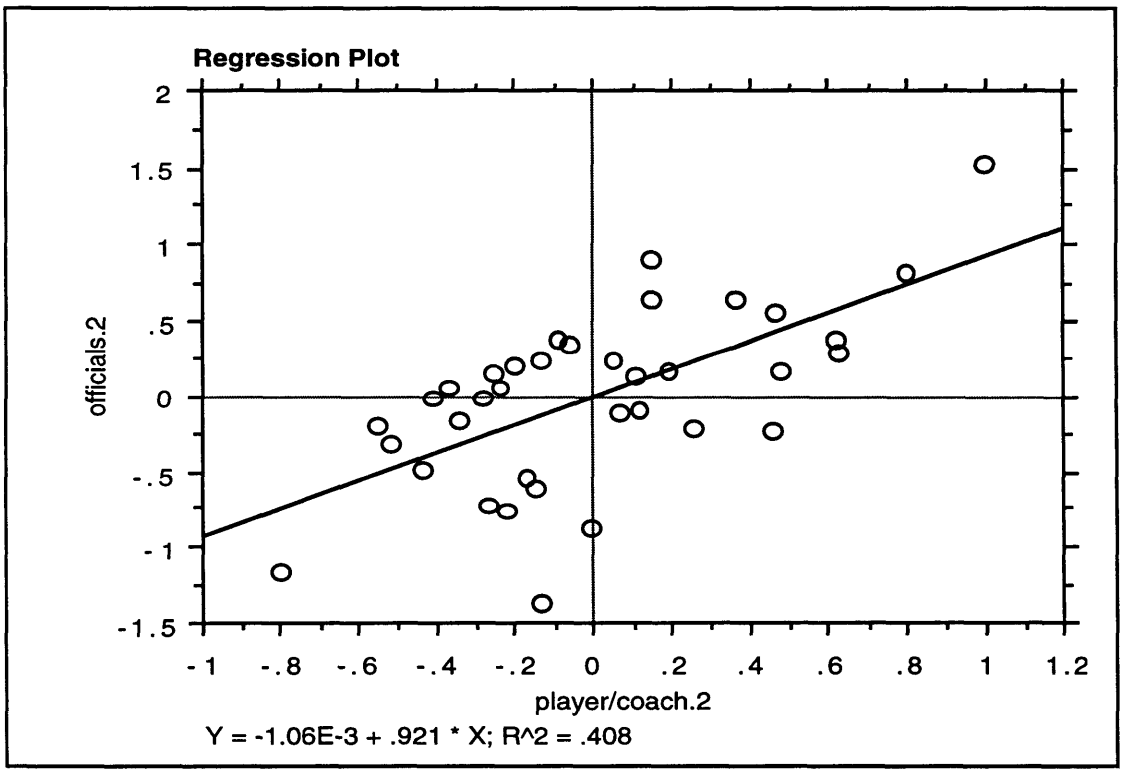

Figure 6.5 Preparedness regression plot

Notwithstanding the non-significant findings from Cook's distance analysis, the identification of competencies contributing distinctly to significant group differences can be determined via rank order comparisons. This technique, similar to that used to distinguish competencies implicated in perspective differences in the previous section, provides a more general mechanism to distinguish specific competencies not detected through Cook's distance.

Rank-Order Comparisons

The specific processes for demonstrating rank order comparisons were detailed in the previous section. However, in review, the rank order of competencies are based on item estimates derived for the preceding Cook's distance analysis. The comparison is achieved by calculating the difference in ranking positions provided by each group, for each competency, within each perspective. A review of ranking comparisons, with respect to the perspectives of *importance* and *preparedness*, is provided in the following discussion.

Importance

Figure 6.6 illustrates *importance* rank-order differentials between groups (complete rank-order inventories are provided in Appendix 24). As can be noted from this figure, a difference in opinion between groups is evident for all competencies except two, i.e., competencies 1.1, *understands and interprets the Laws of Soccer correctly*, and 5.3, *engages in post match activities (e.g., talk to coaches, attend post game functions).*

In the context of this analysis, competencies which recorded the most marked differences (i.e., a differential of seven or more ranking places[1]) are listed in Table 6.12. A diverse range of competencies are evident, however, of interest from this table are competencies ranked high in *importance* (i.e., single figure ranking positions – see Appendix 24). For the competitive group, competencies 2.2 *observes, analyses and correctly interprets incidents*; 2.5 *interprets and discriminates between the severity of fouls*; and, 2.7 *distinguishes between advantage and disadvantage*) fulfil this criteria.

Inspection of these three competencies reveal they are linked closely to the interpretation of rules and incidents. Inasmuch, they effect player conduct, actions, and the immediate nature of the game. Moreover, the application of these competencies in the competitive match environment is subsumed within the cognitive decision-making process. Consequently, their application is variable, and, moreover, the quality of their implementation is difficult to quantify.

---

[1] This criteria is a subjective measure, determined by the investigator for discussion purposes.

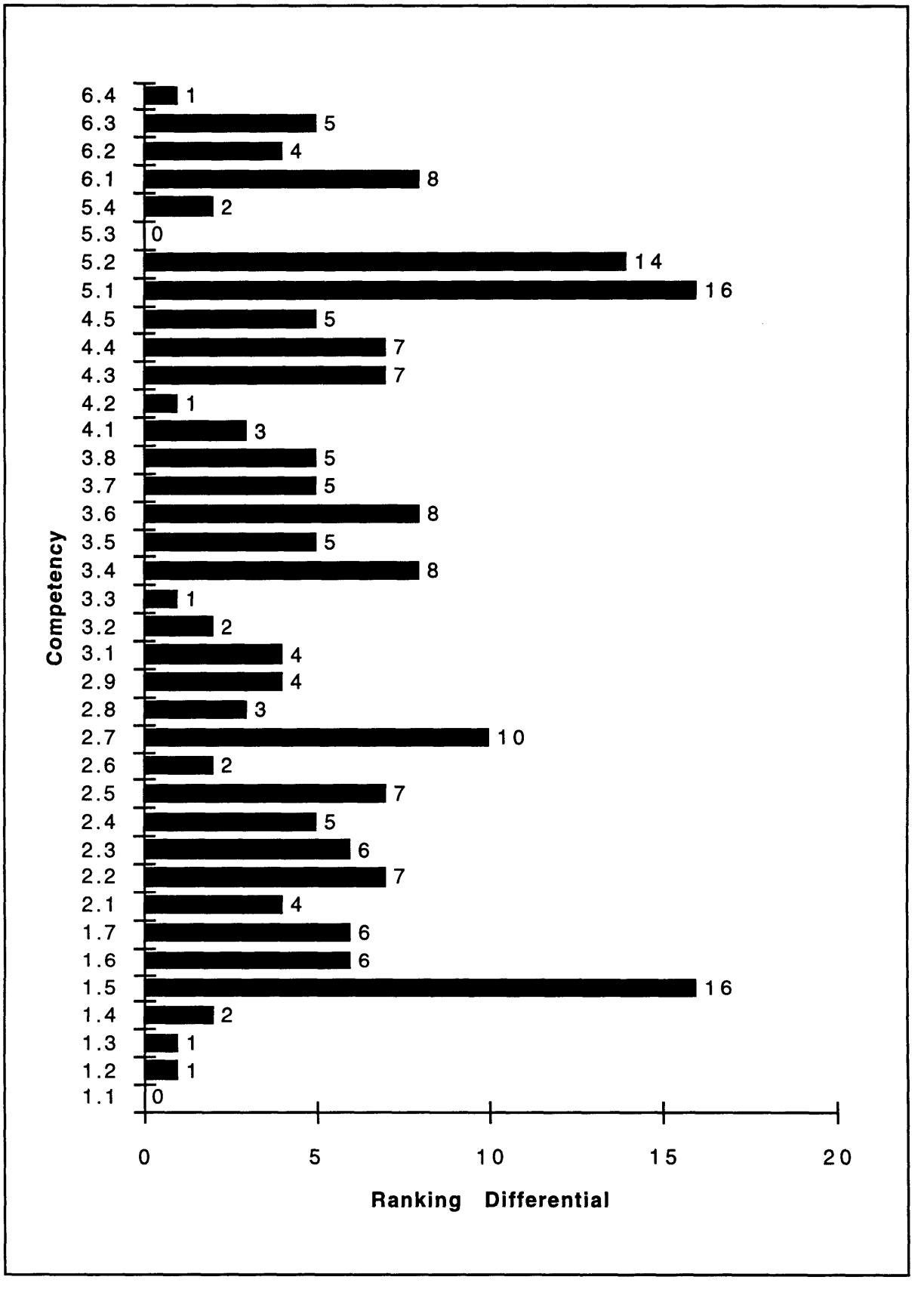

Figure 6.6 Competency ranking differential between groups – Importance perspective

Table 6.12 Marked ranking differential on Importance items by group

Competency	Official's Ranking	Competitors' Ranking	Ranking Differential
1.5 keeps a complete record of the game (e.g., bookings, goals, substitutions)	4	20	16
2.2 observes, analyses and correctly interprets incidents	11	4	7
2.5 interprets and discriminates between the severity of fouls	15	8	7
2.7 distinguishes between advantage and disadvantage	18	8	10
3.4 communicates (verbally and non-verbally) with players on and off the field	24	16	8
3.6 communicates confidently with assistant referees	17	25	8
4.3 manages disputes between players/coaches and match officials	29	22	7
4.4 monitors player behaviour	21	14	7
5.1 prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared)	16	32	16
5.2 undertakes mental preparation for the match (e.g., visualisation)	22	36	14
6.1 maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)	3	11	8

Conversely, the competencies ranked most highly by the officiating group, i.e., competencies 1.5, *keeps a complete record of the game (e.g., bookings, goals, substitutions)* and, 6.1, *maintains required levels of fitness (e.g., fitness benchmarks/standards, hydration, warm up/cool down)* do not possess these characteristics. These three competencies are less subjective to observe, and, therefore, performance is easier to quantify. Moreover, these competencies do not impinge on play immediately, nor do they impinge on the individual actions and conduct of players.

Specifically, the competency of record keeping (1.5) is an administrative and procedural matter which is removed considerably from the decision-making responsibilities of the referee. With respect to competency 6.1 (adherence to fitness standards), its relationship to decision-making is subliminal. It can be argued logically that the quality of the referee's decision is dependent on sound positioning, and, therefore, by implication, the ability of the referee to move into appropriate positions. In short, correct decision-making is linked principally to high fitness levels. However, discussion concerning group differences in the perceived importance of the competency may be exaggerated. Although a ranking differential of eight was recorded (officiating ranking position = 3; competitive ranking position = 11), the relatively high rankings afforded this competency suggests the competency is perceived by both groups to be important.

Other competencies worthy of comment from Table 6.12 are competency 5.1, *prepares well in advance of the match (time of arrival at ground – on time, presentation on arrival, kit prepared)*, and 5.2, *undertakes mental preparation for the match (e.g., visualisation).* Although both competencies recorded ranking differentials (16 and 14, respectively), neither of these two competencies are ranked highly in terms of *importance* by either group. This is particularly so for competency 5.2, which recorded rankings of 22 and 36 from the officiating and competitive groups, respectively. Nonetheless, both competencies do indicate major deviation in opinion, and consequently are seen as contributors to the significant differences in the perception of *importance* between groups.

Preparedness

Figure 6.7 illustrates *preparedness* rank-order differentials between groups (complete rank-order inventories are provided in Appendix 25). As noted from Figure 6.7, only six competencies show a ranking differential of three or less, while 13 competencies recorded a differential of 10 or more places. Two competencies recorded complete agreement between groups, these being competencies 1.5, *keeps a complete record of the game (e.g., bookings, goals, substitutions)*, and 3.2, *communicates decisions with clear hand signals.* In both cases, the competencies were ranked highly (see Appendix 25).

The three items which show the greatest discrepancy between groups are detailed in Table 6.13. With respect to these three competencies, two points emerge. Firstly, the competencies are ranked lowly by officials, and relatively high by the competitive group (particularly the latter two competencies). Secondly, the execution of these competencies occurs away from the competitive aspects of the game.

Table 6.13 Highest ranking differential on Preparedness items by group

Competency	Official's Ranking	Competitors Ranking	Ranking Differential
6.3 undertakes appropriate risk management procedures (e.g., know legal responsibilities, state of the pitch)	28	7	21
6.2 maintains appropriate levels of personal health (e.g., correct diet, weight control)	27	9	18
3.8 communicates with referee's inspectors (e.g., post match discussions, self reflection, constructive criticism)	35	17	18

In a bid to provide more meaningful inquiry, it is prudent to develop a decision rule to guide the analysis. Of interest – particularly for practical purposes – are competencies undertaken in the context of match-play, which are ranked highly by one group (18 or less) and lowly by the other (18 or greater), *and* have a ranking differential between groups of ten or more. Although this decision rule is designed by the investigator and is arbitrary, it does

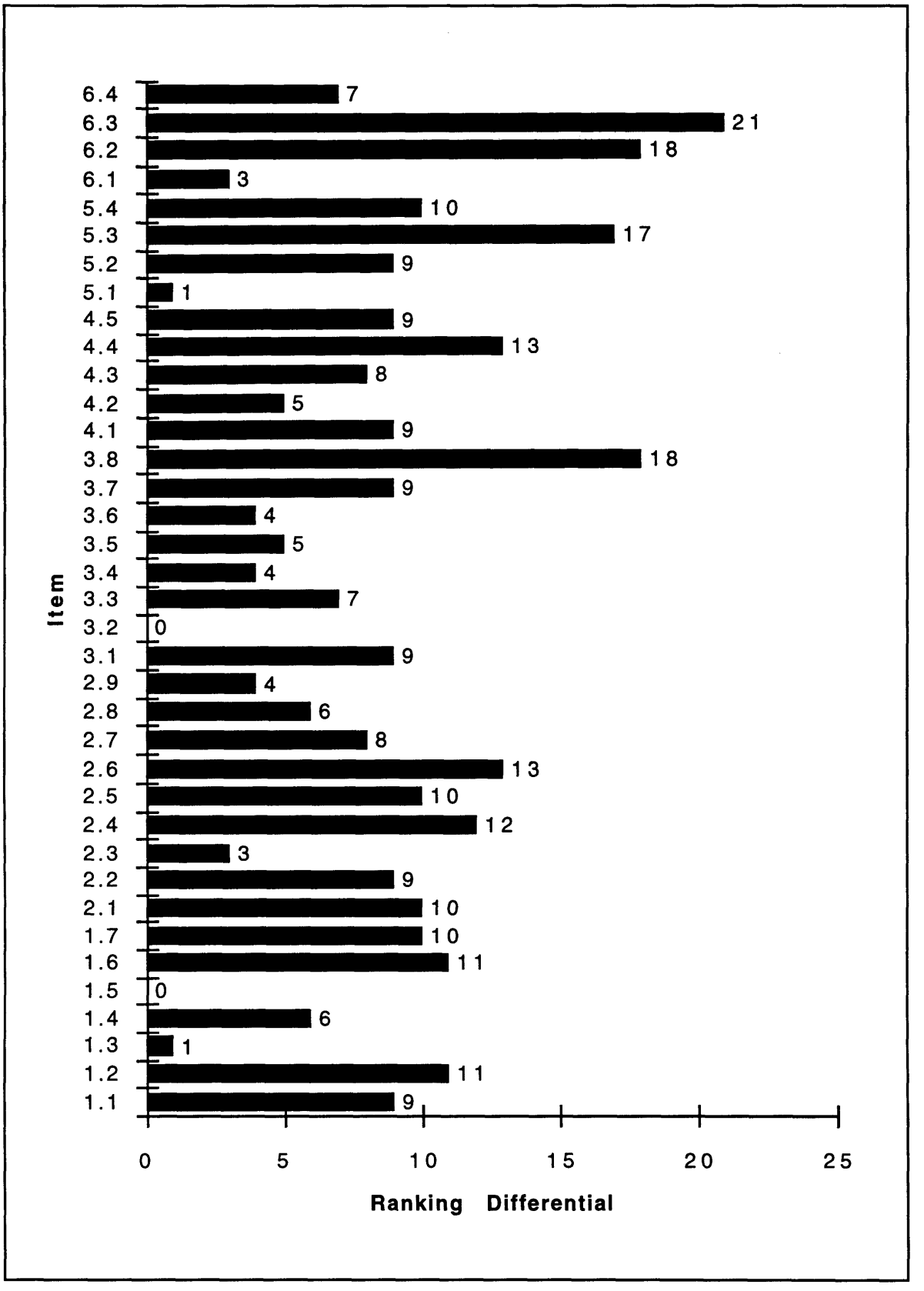

Figure 6.7 Competency ranking differential between groups – Preparation

provide a mechanism for detecting worthy patterns of thought within the findings. Moreover, the rule clarifies which competencies drew marked discrepancies between groups, while ensuring a substantial high-preparedness/low-preparedness dichotomy. Table 6.14 provides ranking details of competencies which meet this decision rule.

Table 6.14 Marked ranking differential on Preparedness items by group

Competency	Official's Ranking	Competitors Ranking	Ranking Differential
1.2 observes incidents and decides on appropriate action (e.g., fouls, player acting, free kick, booking, play-on)	13	24	11
1.6 works as a team with assistant referees	16	27	11
1.7 observes play from the best position	13	23	10
2.4 distinguishes between fair and foul play	7	19	12

Two issues arise from Table 6.14. Firstly, the officials provide higher rankings across all four competencies. Obviously, the officiating group view the preparedness of these competencies more favourably than the competitive group. Reversals in this trend, i.e., competencies ranked highly by the competitive group that also meet the stated criteria, is not evident in the data. Secondly, three of the four competencies (1.2, 1.7, and 2.4) relate to the immediate decision-making ability of the referee.

Overall, the results in this section demonstrate that opinion relating to the *importance* and *preparedness* of refereeing competencies is inconsistent across soccer stakeholders. Differences were shown to be statistically different, although it cannot be concluded statistically which aspects of refereeing performance are most implicated in causing group differences.

Discussion and Implications

Multivariate analysis and variance (MANOVA) showed a significant difference ($p \leq .000$) between groups by perspective. Subsequent univariate analysis (ANOVA) confirmed a significant difference between groups ($p \leq .004$) for both dependent variables (importance, $p \leq .004$; and preparedness, $p \leq .001$). Accordingly, the following hypotheses were rejected:

H 4.1 *There will be no significant difference of opinion about the relative importance of elite refereeing performance criteria among referees assistant referees, referee inspectors, players, and coaches.*

H 4.2 *There will be no significant difference of opinion about the relative preparation of elite refereeing performance criteria among referees, assistant referees, referee inspectors, players, and coaches.*

Follow-up analysis using Cook's distance failed to identify significant outliers from either data set. This indicated that all competencies, within each dependent variable, were implicated in contributing to the significant univariate results. However, by comparing the rank order of competencies, according to group responses within each perspective, obvious differences in group perceptions emerged. In general, this revealed that there was little agreement between groups across most competencies for each dependent variable.

The marked differences recorded for the *importance* data (see Table 6.12) are interesting. A strong emphasis on rule application from the competitive group is not mirrored by the officiating group. Conversely, the officiating groups' predilection for competencies which are easy to quantify, and removed from the 'cut-and-thrust' of play, is an emphasis dismissed by the players and coaches.

Two competencies which did show agreement between groups, i.e., 1.1, *understands and interprets the Laws of Soccer correctly*, and 5.3, *engages in post match activities (e.g., talk to coaches, attend post game functions)*, were opposed on the *importance* continuum. Strong support across groups for competency 1.1 underscores the high overall ranking of the competency (see Table 6.3). The low ranking of competency 5.3 indicated neither group is particularly interested in talking to each other after the game, and is symptomatic of the low rankings afforded other communication competencies.

For the *preparedness* data, three competencies (3.8, 6.2, and 6.3) recorded the largest discrepancy. However, these results should be interpreted with caution. This is because in undertaking these competencies, referees are generally carrying out off-field duties. As such, players and coaches may not be in a position to observe fully the referee's proficiency in these competencies.

In specific reference to this point, communication between referees and referee inspectors (competency 3.8) occurs normally in the changing-room following each match, or in the ensuing week at referee training. Consequently, players and coaches are not privy to such interaction, and thus are possibly not in a strong position to make a valid assessment.

Similarly, the referee's maintenance of personal health (competency 6.2) is also undertaken away from the match environment. Aspects of diet and weight control are personal factors, and are undertaken by the referee either at home, or in the context of their training environment. Perhaps the only concession that can be afforded players and coaches in making judgements on this competency's preparedness is referee appearance. Obviously, if a referee looks overweight, players and coaches can make a direct causal link between the referee's diet/training regime and physical shape. However, this assessment is still subjective, and is susceptible to other extraneous factors, e.g., the referee's somatotype.

Lastly, the referee's responsibility for risk management procedures (competency 6.3) can be examined from within a cognitive context. In undertaking this competency, referees need to make judgements about player safety, and the subsequent liability they [the referee] may incur as a result of breaching their duty of care to players. The development of this awareness occurs through training processes incorporated into referee development, and is enacted before, or after, the playing of a match. Again, this minimises opportunity for players and coaches to view the referee's ability in executing this competency.

Notwithstanding such limitations, the thrust of results with respect to competencies listed in Table 6.14 (i.e., competencies 1.2, 1.7, and 2.4) require comment. These three competencies relate to decision-making aspects of the referee's role. The practical links between the competencies are of marked importance, as the execution of these competencies have an immediate impact on the actions of players, and subsequent conduct and flow of the game. In combination, the competencies require that referees not just observe an incident, but observe play from the best possible position. Subsequently, the referee must decide if the incident represented fair play, and what consequential appropriate action (if any) to take.

Further, the consistently low ranking of these competencies provided by the competitive group, as opposed to the relatively high rankings provided by the officiating group, point to a divergence of opinion concerning fundamental aspects of refereeing. Given that the majority of the competitive group is composed of players (84%), their opinion predominates. This is a salient point, particularly as players can be considered 'clients' of referees' work. As such, their opinions are formed via immediate interaction with the referee, and as a consequence of referee decisions. In combination, these two reasons form a powerful base on which opinions are developed. Accordingly, opinion expressed by the competitive group should be heeded with respect to determining the overall quality of elite soccer refereeing. Clearly, the recipients of refereeing decisions feel greater preparation in these skills is required.

CONCLUSION

This chapter addressed a number of research questions and hypotheses. These ranged from determining: which competencies were important; the level of competency preparedness; if significant differences existed between the perspectives of *importance* and *preparedness*; and, if perspective differences were consistent across groups. A range of non-parametric and parametric techniques were used to address these issues, including frequency percentages, Rasch analysis, multivariate and univariate analyses, Cook's distance, and ranking differentials.

Results indicated that competencies are not of equal importance nor preparedness. Indeed, it became evident that the *importance* of competencies, and their *preparedness*, were dissimilar.

Using a two-tail *t* test for paired samples, this difference was shown to be significant (df=36, $p \leq .001$). It was found also that group differences existed within the perspectives of *importance* and *preparedness*. Using a combination of multivariate and univariate analysis, these differences were found to be significant (*importance*: $p \leq .004$; *preparedness*: $p \leq .001$).

Parametric analysis did not clarify which competencies were implicated significantly in causing these perspective and group differences. However, non-parametric rank-order comparisons were able to provide some insight into these differences. With these findings in mind, the following chapter explores the responses of respondents to the open question contained in Section 3 of the questionnaire. In doing so, it attempts to provide detail as to why discrepancies in perspectives and groups were evident.