

Northumbria Research Link

Citation: Howat, Douglas James (1998) Gender stereotypes and children's attitudes towards males and females in sport. Doctoral thesis, University of Northumbria at Newcastle.

This version was downloaded from Northumbria Research Link:
<https://nrl.northumbria.ac.uk/id/eprint/15696/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

Some theses deposited to NRL up to and including 2006 were digitised by the British Library and made available online through the [EThOS e-thesis online service](#). These records were added to NRL to maintain a central record of the University's research theses, as well as still appearing through the British Library's service. For more information about Northumbria University research theses, please visit [University Library Online](#).



**Northumbria
University**
NEWCASTLE



University**Library**

GENDER STEREOTYPES AND CHILDREN'S ATTITUDES TOWARDS MALES AND FEMALES IN SPORT

DOUGLAS JAMES HOWAT

**A thesis submitted in partial fulfilment
of the requirements of the
University of Northumbria at Newcastle
for the degree of Doctor of Philosophy.**

September 1998

CONTAINS DISKETTE

-UNABLE TO COPY-

**CONTACT
UNIVERSITY IF YOU
WISH TO SEE THIS
MATERIAL**

ABSTRACT

Along with other researchers, the Sports Council (1994b) has highlighted the negative influence of gender stereotypes on female involvement in sport. The aim of the present research was to investigate children's stereotypes of, and attitudes towards, males and females in sport. A series of five related studies was conducted. Four of these studies were based on a disguised attitude measure, the Goldberg paradigm (Goldberg, 1968). Participants were presented with a questionnaire that included a brief description of a character whose sex and sport participation, unbeknownst to the participants, were varied across the conditions. Participants were asked to rate the characters on a variety of dependent measures that were designed to allow differential evaluations of the male and female characters. In Studies One and Two, evidence of gender stereotyping and sport stereotyping was apparent but, contrary to expectations, the data did not indicate that the children negatively evaluated the female sports participant. In Study Three, a sample of parents was presented with the same questionnaire as used in Study Two. Their responses echoed those of the children in the first two studies. Study Four employed semi-structured interviews, the transcripts of which were analysed qualitatively, to directly assess children's attitudes. In contrast to the first two studies, the children demonstrated strongly-held negative attitudes towards males and females taking part in sports that were not considered appropriate for their sex. Evidence was found to suggest that participants were, in line with social judgeability theory (Leyens et al., 1993), respecting social norms when making judgements of other people. Study Five attempted to explicate the results of the previous four studies by again employing a questionnaire based on the Goldberg paradigm. However, Study Five provided no evidence to support the predictions of social judgeability theory. While taken together the results of the five studies were inconclusive, it is argued that social judgeability theory remains a plausible explanation of the findings and that future research should continue to examine this perspective. However, it is argued on theoretical and methodological grounds that future researchers should be wary of using the Goldberg paradigm and that interviews may be a more fruitful way to assess gender stereotypes in sport.

Table of Contents

ABSTRACT.....	2
ACKNOWLEDGEMENTS.....	8
CHAPTER 1: SPORT, GENDER AND THE INFLUENCE OF STEREOTYPES	9
Background to the Research.....	9
The Importance of Sport.....	10
Participation Patterns in Sport: The Nature of the Problem.....	11
Female Employment in Sport.....	15
Explanations of Female Participation in Sport.....	18
Sex Differences Approach	19
Gender-Role Orientation.....	22
Role Conflict.....	24
Motivational Orientation.....	25
Self-efficacy.....	26
Attributions.....	28
Socialisation.....	29
Gender Stereotypes : The Common Link.....	30
Summary.....	34
Stereotypes and Stereotyping.....	35
Definitions.....	35
Early Research on Stereotypes.....	38
Gender Stereotypes.....	40
Content of Gender Stereotypes.....	40
Acquisition Of Gender Stereotypes.....	42
General Theories of Stereotyping.....	44
Psychodynamic Theories.....	44
Realistic Conflict Theory.....	46
Social Learning	47
Self-Categorisation Theory and Social Identity Theory.....	48
Impression Formation.....	53

Social Judgeability Theory.....	56
The Best of Both Worlds: Combining Social Identity Theory	
and Impression Formation Models.....	56
Empirical Support for Social Judgeability Theory.....	58
The Present Research.....	62
 CHAPTER TWO: THE ASSESSMENT AND MEASUREMENT OF	
STEREOTYPES AND ATTITUDES.....	64
Multiple Research Methods	64
Multiple Methods in Social Science Research.....	65
Multiple Methods in Stereotyping Research.....	66
Summary.....	68
Combining Quantitative and Qualitative Methodologies.....	69
The Development of Stereotype and Attitude Measures.....	73
Direct Methods of Stereotype and Attitude Assessment.....	74
Rating Scales.....	76
Free-response Measures.....	81
Interviews.....	82
Interviews in Attitude and Stereotyping Research.....	85
Indirect Methods of Stereotype and Attitude Assessment.....	87
The Link Between Attitudes and Behaviour	87
The Lost Letter Technique.....	91
The Wrong Number Technique.....	92
The Goldberg Paradigm.....	93
Problems with Indirect Methods	94
Summary.....	95
The Present Research.....	96
Studies One, Two, Three and Five:	
The Goldberg Paradigm Studies.....	96
Study Four: The Semi-structured Interview Study.....	99
Summary.....	100

CHAPTER THREE: THE GOLDBERG PARADIGM (MK 1) :

STEVEN AND SUSAN PLAY THE VIOLIN AND GO SWIMMING..... 102

Introduction to Study One..... 102

Method..... 106

 Design..... 106

 Questionnaire Design..... 106

 Participants..... 120

 Procedure..... 120

Results..... 122

Discussion..... 128

 Summary..... 136

 Conclusions..... 137

CHAPTER FOUR: THE GOLDBERG PARADIGM (MK II): REBECCA

AND DANIEL PLAY RUGBY AND GO SWIMMING..... 139

Introduction to Study Two..... 139

Method..... 144

 Design 144

 Questionnaire Design..... 144

 Participants..... 156

 Procedure..... 156

Results..... 157

Discussion..... 168

 Summary..... 172

CHAPTER FIVE: THE GOLDBERG PARADIGM (Mk III):

WHAT DO PARENTS THINK OF REBECCA AND DANIEL

PLAYING RUGBY AND GOING SWIMMING?..... 174

Introduction to Study Three..... 174

Method..... 176

 Design..... 176

 Questionnaire Design..... 176

Participants.....	177
Procedure.....	178
Results.....	178
Discussion.....	191
Summary.....	196
CHAPTER SIX: THE INTERVIEW STUDY: A DIRECT ASSESSMENT	
OF CHILDREN'S ATTITUDES AND STEREOTYPES.....	198
Introduction To Study Four.....	198
Method.....	202
Design	202
Participants.....	202
Pre-pilot Study and Pilot study.....	202
The Interview Schedule.....	203
Procedure.....	205
Data Analysis.....	207
Results and Analysis.....	209
Overview: Sports Participation.....	209
Attitudes towards females in sport.....	210
Attitudes towards males in sport.....	213
Social Norms and the Expression of Attitudes.....	215
Discussion.....	218
Conclusions.....	220
CHAPTER 7: THE GOLDBERG PARADIGM (MK IV): REBECCA	
AND DANIEL PLAY NETBALL, RUGBY AND GO SWIMMING.....	221
Introduction to Study Five.....	221
Manipulation of Instructions.....	222
Manipulation of Sport.....	223
The Dependent Measures.....	224
Method.....	226
Design.....	226

Questionnaire Design.....	226
Participants.....	233
Procedure.....	233
Results.....	236
Discussion.....	249
Conclusions.....	253
CHAPTER EIGHT: GENERAL DISCUSSION.....	253
Implications for Future Research.....	261
Conclusions.....	263
REFERENCES.....	265
APPENDIX A: Sample questionnaire from Study One.....	283
APPENDIX B: Sample questionnaire from Study Two.....	290
APPENDIX C: Sample questionnaire from Study Five.....	299
ON DISC ON INSIDE BACK COVER:	
APPENDIX D: Anova Summary Tables for Study One	
Filename: Studyone.lst (SPSS output format).....	on disc
APPENDIX E: Anova Summary Tables for Study Two	
Filename: Studytwo.lst (SPSS output format).....	on disc
APPENDIX F: Anova Summary Tables for Study Three	
Filename: Studythr.lst (SPSS output format).....	on disc
APPENDIX G: Anova Summary Tables for Study Five	
Filename: Studyfve.lst (SPSS output format).....	on disc
APPENDIX H: Study Five data set	
Filename: Studyfve.sav (SPSS datafile format).....	on disc

Acknowledgements

I would like to thank Dr. Sandy Wolfson and Dr. Lesley Fishwick for their enthusiastic and helpful supervision of my post-graduate studies. Since starting work as a lecturer at Coventry University I have come to realise that I learned much more from them than this thesis demonstrates.

I am exceptionally grateful to all the teachers and pupils at the schools in Newcastle and in Cambridgeshire that participated in my research. I would like to give special thanks to Malcolm Jarvie and his colleagues for the amount of time and effort that they spent arranging their classes around my data collection.

Chris Dracup, from the University of Northumbria, and Chris Wright and Steve Jones, from Coventry University, all gave me useful advice and practical help at various points during the completion of this thesis, and I would like to express my sincere gratitude. My colleagues in the postgraduate research centre at North St. East also deserve thanks. While their help was somewhat less practical in nature their support has been much appreciated. I also appreciate greatly the support of my family, and especially my parents, who have continued to encourage me throughout my education.

Finally, I wish to express my love and thanks to Hazel.

CHAPTER 1: SPORT, GENDER AND THE INFLUENCE OF STEREOTYPES

Sport is a cultural activity which, practised fairly and equitably, enriches society and friendship between nations. Sport is an activity which offers the individual the opportunity of self-knowledge, self-expression and fulfilment; personal achievement, skill acquisition and demonstration of ability; social interaction, enjoyment, good health and well-being. Sport promotes involvement, integration and responsibility in society and contributes to the development of the community.

The Brighton Declaration (Sports Council, 1994a)

Background to the Research

As the extract from the Brighton Declaration above demonstrates, sport is an exceptionally important and powerful institution for individuals as well as for nations. However, as the Brighton Declaration (Sports Council, 1994a) states, female participation in sport (regardless of where in the world) is always less than that of males. Gender stereotypes have been identified as amongst the barriers preventing females from participating fully in sport (Sports Council, 1994b). The present research aims to contribute to the body of knowledge concerning female participation in sport by investigating the stereotypes that children, in particular, hold of males and females in sport.

The first part of this chapter will describe the patterns of male and female involvement in sport and highlight why these patterns are of concern. The second part of this chapter will discuss some of the major theories which have attempted to explain why these participation patterns occur. Within nearly all of these theories gender stereotypes have been identified as a primary cause of discrimination. However, it

will be demonstrated that there has been a lack of empirical research on the nature and content of these stereotypes and the process of stereotyping within a sporting context. The third section of this chapter, therefore, will discuss theories of stereotyping and explain the need for gender stereotypes in sport to be investigated more thoroughly if the aim of providing males and females with equal opportunities is to be achieved.

The Importance of Sport

In Britain today, sport is a multi-million pound industry providing employment and business opportunities to many people. For example, in 1991 it was estimated that £9.75bn was spent by British consumers on sport and related services (Henley Centre cited in Sports Council, 1992). To put this figure in perspective, a study in 1990 estimated that the 'output' from the sport sector yielded 1.7% of the gross national product. This represents a substantial contribution to the economy of the U.K. (Sports Council, 1992). Furthermore, the Department of Employment estimated that 467,000 people were employed in sport-related occupations in the United Kingdom. This is a significant proportion of the 25.6 million total workforce (Sports Council, 1992). While the British economy is complex and multifaceted it is clear that sport has a significant contribution to make. Thus the importance of sport to Britain's economy must not be underestimated as policies which affect sport will have considerable influence on other sectors.

Financial and economic factors withstanding, the sport sector is also important because it is through sport that many people become active and can gain the benefits associated with regular physical exercise. The Allied Dunbar National Fitness Survey (Sports Council and Health Education Authority, 1992) lists the benefits of sport as being a reduced risk of coronary heart disease, relief of mild hypertension, increased stamina, prevention of osteoporosis, management of some forms of diabetes, maintenance of muscle strength and joint flexibility, management of body weight, reduced stress, and enhanced mood and self-esteem.

Given the medical benefits and economic significance of sport there is great concern for any group or section of society which is excluded from sport. The 'Sport for All' principle is enshrined in the European Sports Conference Charter which states that "Sport is an inalienable right of every person" (Sports Council, 1992). Despite this declaration of 'right', as will be shown, sport is not open to all groups within society and this has led to many programmes and projects aimed at involving more people, from a wider range of backgrounds, in sport. In particular females, members of ethnic minorities and people with disabilities have been identified as groups under-represented in sport (Sports Council, 1994b). The focus of the current research is female participation in sport but it is not intended that this area should be seen as more or less important than other areas of equal opportunities work within sport, and it is acknowledged that this is only one important aspect of equal opportunities work amongst many.

Participation Patterns in Sport: The Nature of the Problem

It is overly simplistic to suggest merely that females are under-represented in sport. Participation patterns are complex and continually changing. As the figures presented later in this chapter will show, there are some areas of sport where females comprise the majority of participants (indoor sports in particular) and these sports could be cited as evidence that females are not discriminated against. However, a more detailed examination of participation and employment figures is needed to demonstrate that while female sport participation is increasing, females are still excluded from many areas of sport while being given almost exclusive access to other areas.

It is clear that more females are taking part in sport than ever before. This is true for elite levels as well as for 'participation' levels. For instance, during the twentieth century the number of females taking part in the Olympic Games has risen steadily from 0.9% of the total number of participants at the Paris games of 1900, to 40% at the Barcelona games of 1992 (Hargreaves, 1994). This pattern is also reflected in

non-elite sporting contexts. Figures from the General Household Survey carried out in the U.K. show that during the late 1980s the number of females taking part in any sporting activity (including activities such as walking) in the four weeks prior to the survey, increased from 53% to 57% of the adult female population. While this increase contributed to the overall rise in adult sport participation from 60.7% to 64.5% of the total population, the percentage of adult males taking part in sport (73% of adult male population) is still considerably higher than that of females (Sports Council, 1992).

While the gap between male and female participation rates is certainly narrowing, the overall figures mask the underlying trends. For instance, much of the increase in female participation in sport can be explained by the increase in popularity of indoor sports and, in particular, aerobics. Females constitute 60% of those people who take part in indoor sports only. According to the British Sports Council this reflects the success of promotions which depict sport as a healthy activity and the general increase in knowledge about the causes of ill-health and the remedies provided by regular physical activity (Sports Council, 1994b).

While data from surveys such as the General Household Survey are undeniably useful there is a problem with the way in which 'sport participant' is defined. The General Household Survey recorded whether an individual had or had not taken part in any sport during the four week period prior to the survey. Anyone who had taken part in any sport during the four week period was defined as a 'participant' regardless of the number of times during that period the sport was played, or the level or intensity of the sporting activity. This creates a problem by placing people with very different levels of sporting involvement in the same category. For instance, a professional footballer, training 4 times a week, and playing 2 competitive matches, would be classed as a sports participant, as would a person who took their dog for a 2 mile walk once in four weeks.

The Allied Dunbar Fitness Survey, however, did investigate levels of physical activity in a study of English adults (Sports Council and Health Education Authority, 1992).

The findings showed that a great deal of vigorous physical activity is accounted for by sporting activities, so this survey's analysis of intensity and frequency of physical activity can also provide more details of the differences between male and female involvement in sport. This survey classified people into 6 'activity levels' ranging from 'Activity Level 0' (no activities of 20 minutes duration in the previous four weeks) to 'Activity Level 5' (12 or more occasions of vigorous activity of 20 minutes duration in previous four weeks). Each level incorporated elements of duration, frequency and intensity. The activity levels for males and females (age range 16 to 74 years) are summarised in Table 1.1. While there is not much difference for males and females on the lowest two activity levels (Level 0 17% males, 16% females; Level 1 16% males, 18% females), males considerably outnumber females on the higher levels (Level 5 14% males, 4% females). Taking into account the figures from the General Household Survey, it appears that females may be taking part in sport in greater numbers but the levels of physical activity remain comparatively low.

Activity Level	Men %	Female %
Level 5 (most active)	14	4
Level 4	12	10
Level 3	23	27
Level 2	18	25
Level 1	16	18
Level 0 (least active)	17	16
Total	100	100

Table 1.1 Allied Dunbar Fitness Survey : Activity Levels for males and females
(Sports Council and Health Education Authority, 1992)

Although the Allied Dunbar Survey did cover level and duration of activity, unlike the General Household Survey, both surveys only covered adults' sport and physical activity. National surveys which have looked at children's sports participation have been limited (Sports Council, 1988, 1993). This is a problem for any attempts to explain female participation in sport because, as will be shown in later sections, people's participation in sport as adults is greatly influenced by their experience of sport as children.

One study which did investigate children's sports participation (Hendry, 1993) found that boys played more sports than girls, and they took part in sport more often. Furthermore, boys were more likely to be members of a sports club or to have played for an organised team, suggesting that gender differences in sports participation amongst children are similar to differences found amongst adults.

As noted previously, programmes and promotions to encourage more people to take part in regular physical activity have been particularly successful at attracting females into indoor sports such as swimming and aerobics, and as a result the figures for overall sport participation have increased for females. However, there are still areas of sport that remain largely impenetrable to females. For instance, females comprise only 6.7% of participants in outdoor team sports (Sports Council, 1992). It is interesting to note that Iain Sproat, the Junior Heritage Minister responsible for sport, while extolling the benefits of sport, belittled (perhaps inadvertently) the activities that females are becoming most involved in:

Sport is so important because it affects the whole character of a generation, let alone its health. And when I say sport I do not mean aerobics, stepping up and down bars or countryside rambles. What I mean is properly organised team games..... the traditional games of this country: soccer, cricket, hockey, rugger, netball. (Times Education Supplement, July 12th 1996)

It appears, therefore, that while overall participation figures show increased female sport involvement, the sports in which females dominate, such as aerobics, are devalued as not 'proper' sports.

When introducing policies that affect sport, the U.K. government should take into account the influence that legislative intervention has had on sport participation in the U.S.A. during the last 25 years. Although comparisons with the U.S.A. are not unproblematic, it is of great concern, that even with the help of unique legislation (Title IX), the increase in female sports participation in the U.S.A. has slowed in recent years. For instance, in 1971, only 7% of interscholastic athletes were female. This figure increased to 34% in 1981. However, in the next 12 years to 1993, female participation only increased by a further 3% (Houlihan, 1997). Bucher and Krotee (1998) suggest that while more females are taking part in sport, it is not yet a 'level playing field'.

There is some evidence that a similar trend is occurring in Britain. While the number of females attending Premier League football matches has increased, only 12% of fans are female (Tester, 1998). Of those men that attend Premier League football matches, 55.8% have also played football as adults, compared with only 5.5% of female fans (Tester, 1998). So while these figures represent an increase in female participation in football, both as players and spectators, female participation remains considerably lower than that of males.

Female Employment in Sport

There is more to developing female involvement in sport than increasing the number of females taking part in specific activities. Houlihan (1997) reports that in the U.K., as in the U.S.A. and Australia, while overall sports participation by females is increasing, participation in administration and coaching roles has decreased in the last 20 years. Sport-related employment, such as coaching, physical education and sport journalism, is of great importance, as individuals doing these types of jobs have the

ability and power to influence the development of sport and female participation in sport. (Sports Council, 1994b).

Table 1.2 shows the percentages of male and female coaches at the last 3 summer Olympic Games, in comparison to the percentages of male and female athletes. It can be seen that the proportion of female athletes is increasing though, as pointed out earlier, equity has not been achieved. This disparity however, is much greater when the percentages of coaches are considered. In 1992 only 8% of the coaches at the Olympic games were female despite the fact that 39% of the athletes were female. A similar pattern can be seen in the preceding games with a tiny proportion of female coaches in comparison to the number of female participants.

	% of athletes		% of coaches	
	male	female	male	female
1984	68	32	96	4
1988	64	36	90	10
1992	61	39	92	8

Table 1.2 Percentages of male and female athletes and coaches in the Summer Olympic Games (Sports Council, 1994b)

The small number of female coaches may be expected in sports such as football or rugby where, traditionally, very few females have played the sport and gained the necessary experience prior to becoming a coach. It might also be expected that as greater numbers of females take part in these sports there will be an increase in the numbers of females 'graduating' from the game into coaching positions. However, there are sports which have a large number of female participants and which still have very few female coaches. Females make up approximately 44% of the estimated 2 million badminton players in England and yet in 1990 only 20% of county coaches were female, and there were no female national coaches (Sports Council, 1994b).

Gymnastics is a sport which has separate male and female disciplines, and in 1976 only men coached the male discipline and only females coached the female discipline. As Table 1.3 shows, by 1990 however, 52% of the coaches in the female discipline were male, while only 13% of the male discipline coaches were female (Sports Council, 1994b).

	% Coaching male discipline		% Coaching female discipline	
	Male	Female	Male	Female
1976	100	0	0	100
1990	87	13	52	48

Table 1.3 Percentages of male and female gymnastics coaches (Sports Council, 1994b)

Figures show that females are under-represented in other sport-related employment. A survey carried out by the Northern Region Sports Council (1992) investigated the number of males and females in a variety of jobs and again found that females are under-represented in most areas and over-represented in others. In particular it was found that males outnumbered females in sports journalism, headships of school P.E. departments and sports science. One exception was that females outnumbered males in physiotherapy (Sports Council Northern Region, 1992).

Taken together, these data show that females are under-represented in most areas of sport employment, and over-represented in others. There is evidence to suggest that female participation is particularly concentrated in low-status sectors within sport and that the 'glass ceiling', the invisible barrier preventing females achieving high-status positions, is "lower and more impervious than in other domains" (Gill, 1995).

Because there are few females in leadership positions and sports-related employment their ability to influence decisions about sport is minimised (Sports Council Northern Region, 1992).

Overall these data can be summarised as follows :

1. Female participation in sport is increasing
2. The increase in female participation in sport is due largely to relatively few sports, particularly those sports which emphasise health and fitness.
3. Even in sports with large numbers of female participants there are few female coaches.
4. There are few females in influential leadership roles within sport (e.g. education, media)

This pattern of female sport participation has led the British Sports Council to endorse the aim of the European Sports Conference "to increase the involvement of women in sport at all levels and in all functions and roles" (Sports Council, 1994b). Statistics may be illustrative but general trend data do not explain why these patterns of sport involvement occur and they do not provide any solutions. Hence the Sports Council has endorsed the further objective "to increase scientific knowledge about women and sport", and thus continue and build on work over the past 25 years by social science researchers, who have attempted to explain patterns of female participation in sport (Sports Council, 1994b).

Explanations of Female Participation in Sport

As the following review will show, there have been several different approaches which have attempted to explain the patterns of female participation in sport. These approaches include explanations based on attribution theory, self-efficacy,

socialisation, and gender-role orientation amongst others. One problem that these approaches suffer from is a lack of integration with each other (Gill, 1995; Plaisted, 1995). However, one factor that the more recent approaches have in common, in line with the Sports Council (1992), is that they identify gender stereotypes as amongst the barriers preventing females from participating in sport. The aim of the following review is not to assess the relative merits of the different approaches, or to identify the most convincing or promising explanation of female participation patterns in sport. The present research has been carried out with the underlying philosophy that the approaches concentrating on one variable or construct have been too simplistic. Instead, successful explanations will be complex, include many variables, and take into account the role of context on social behaviour (Weiss & Glenn, 1992; Plaisted, 1995). The following review is presented to demonstrate that any explanation of male and female sport involvement must take into account gender stereotypes.

Sex Differences Approach

One of the first explanations put forward to explain female under-representation in sport was the sex differences approach. The main assumption of the sex differences approach was that differences in psychological abilities stemmed from biological differences between males and females (Gill, 1995).

In a comprehensive review of psychological abilities and characteristics, Maccoby and Jacklin (1974) suggested that males might be more aggressive than females, have better mathematical abilities and have better visuo-spatial abilities. The proposal put forward was that males were superior at sport because of these differences in abilities which are seen as central to success at many sports. This proposal was widely accepted despite the fact that Maccoby and Jacklin only put forward the *possibility* that these differences might exist (Gill, 1995). Indeed, subsequent meta-analytical research has found that differences in these abilities are small, inconsistent and highly dependent on context and the nature of the task (Eagly, 1987; Hyde, 1990).

Hyde (1990) rejects the suggestion that these small differences can be used to explain sex-segregation in certain occupations. For example, it may be the case that engineering depends very heavily (though not exclusively) on spatial ability, such that only those in the top, say 10%, of the population in terms of spatial ability are able to become engineers. As males demonstrate a slight advantage over females on tests of spatial ability it would be expected that a larger proportion of the top 10% would be males. However, Hyde estimates that the meta-analytical effect size of 0.40 (i.e. a slight male superiority) for spatial abilities should result in 33% of engineers being female and not the observed 5%. It is not simply the case that engineering jobs are filled by those people who demonstrate a superior ability in spatial tasks. Similarly in sport, while sex-differences might be used to explain why there are fewer female participants, this does not explain why the differences in male and female participation are so great.

It is also possible to challenge the sex differences approach on more fundamental grounds. For instance, while it could be argued that certain occupations require people with exceptional skills, and that because males are (possibly) better than females on the required skills, there are going to be more males than females in those occupations. The converse is also true for jobs which require skills in which females outscore males. However, sport is played at many different levels of ability and competitiveness. To take part in sport does not necessarily require exceptional skills or abilities. The sex differences approach might be able to explain why the men's world record for the 100m. is better than the female world record, but it is unable to explain why fewer females take part in sport at all levels.

A further problem with the sex differences approach is that it does not explain why females do not play sport with other females. Even if men can be shown to have a biological advantage over females, there are very few sports where males and females compete against each other. Any differences between the sexes should be irrelevant because males and females usually have separate disciplines. For instance, there is no logic in the argument that says that females do not take part in sport because the average male is stronger than the average female. A large percentage of females are

still stronger than the average male (i.e. within sex differences are greater than between sex differences), so there is nothing in the nature of sports *per se* which would prevent females from taking part. This is the same regardless of the characteristic being considered, whether it is strength, aggression or visuo-spatial abilities.

It is interesting to note that in boxing, where height and weight confer considerable advantage, there is no question that smaller and shorter men should not be allowed to compete. Instead there are many strict weight divisions to ensure that boxers only compete against those of similar physique.

Another failure of the sex differences approach is the question of causality. As Hyde (1990) and Dyer (1982) point out, the critical issue of whether any of these differences in mental abilities are biological in origin or are the result of socialisation remains unanswered. That is, any measurable differences in the characteristics of males and females could be caused by the fact that males play more sport than females. Males might be more aggressive *because* they play sports, and males might be better at visuo-spatial tasks *because* of their sport involvement. Even if sport involvement is taken into account, the activities which are expected of males and females from birth are so different that it is impossible to make firm conclusions about the causes of sex differences (Douglas and Miller, 1977).

To summarise, the sex differences approach lost much support during the 1970s for three main reasons. Firstly, many of the proposed sex differences have not been found to be as consistent or as large as proposed (Dyer, 1982). Secondly, the sex differences approach has been unable to demonstrate whether any differences in males' and females' psychological abilities result from or cause the different patterns of sport involvement (Dyer, 1982). Thirdly, most sections of sport have separate competitions for males and females so even if there are differences between the sexes in abilities related to sport this does not explain why so few females take part in sport.

One of the major impacts of this work however, was to reinforce traditional stereotypes of males and females (Plaisted, 1995). Despite inconsistent and inconclusive results, the search for sex differences has led to an increased belief that males and females are fundamentally different along dimensions that are important for sport. As will be seen in later sections, the 'kernel of truth' debate, which asks whether stereotypes are exaggerated representations of real differences, renders intractable problems. That is, the question of causality can never be resolved. The sex differences approach tried unsuccessfully to establish real differences. What is of more interest is people's *perceptions* of sex differences i.e. their gender stereotypes (Gill, 1995). It seems that people believe that males and females are biologically different and this is used to justify and explain low female participation in sport, while ignoring within-sex differences. It is important to establish what beliefs people have about males and females in sport to be able to understand this process of justification and the ultimate exclusion of females from sport.

Gender-Role Orientation

The failure of the sex differences approach led researchers to look instead at personality and individual differences without the assumption that the differences were biologically determined. The problem of causality that faced the sex differences approach was avoided and the question of whether sex differences were biological or cultural in origin was largely unaddressed.

One of the first personality constructs to be applied to sport was that of gender-role orientation. The 'gender role orientation' of an individual is not necessarily linked to biological sex. Using the Bem Sex Role Inventory (BSRI), individuals rate a list of traits indicating how true they are of themselves. The traits include 'masculine' and 'feminine' items, previously judged by 100 American undergraduates to be typically desirable for one sex more than the other. Unlike traditional views of 'masculinity' and 'femininity', the BSRI does not conceptualise these two constructs as opposite poles on the same scale. Instead 'masculinity' and 'femininity' are seen as independent

scales. Thus an individual of either sex can score highly on both the masculine and feminine scales. Individuals who score highly on both are classed as "androgynous" and are considered to be more flexible and adaptable than other individuals (Bem, 1974).

The general finding of studies employing the BSRI is that sports participants and female athletes are more likely to be masculine or androgynous than female non-participants (Gill, 1992). However, these findings are not particularly surprising. Sport is very much a patriarchal institution and has developed to emphasise those traits generally associated with males. It is not informative to find that females who take part in sport show 'masculine' characteristics. The fact that 'competitive' is a masculine trait on the BSRI means that any female sport participant is going to score highly on the masculine scale (Gill, 1992). Another major criticism of this research is the use of terms such as 'androgyny' as if they were real constructs without questioning their validity (Hall, 1988). Plaisted (1995) suggests that even the use of terms such as 'masculinity' and 'femininity' seriously limits this line of research, and reifies "destructive gender stereotypes". That is, the more that researchers investigated 'androgyny' the more they came to believe in the concept even though evidence for its existence was mixed.

Other problems include the generalisability in the use of the BSRI. The original 'desirable' traits were selected by American college students. It is questionable whether the BSRI can be used on other cultures and with other age groups. There is good reason to believe that sex-role stereotypes have changed since Bem's original work in the 1970s. While Bem's work generated much data regarding sex-role stereotypes and how individuals see themselves it ignored the issue of context. That is, people may view themselves (and others) in different ways depending on the situation (Plaisted, 1995). Furthermore, there is more to knowledge about gender than the traits associated with males and females. Gender stereotypes also include beliefs about physical appearance, expected behaviours and sexuality amongst others. Concentrating on desirable traits seriously limited the contribution of the BSRI to our understanding of gender stereotypes in sport.

More recently, the traits labelled previously as 'masculine' and 'feminine' have been renamed 'instrumental' and 'expressive' (Spence and Helmreich, 1978). This however still resorts to a 'stereotypical dichotomy' and Gill (1992) suggests that this is still too simplistic a way to consider gender. Rather than seeing gender as a characteristic of individuals, researchers should view gender as a 'social relational process' (Plaisted, 1995; Gill, 1995). It is not enough to know that males and females differ or are believed to differ; the context in which people are being viewed must also be considered, as well as the large within sex differences.

Bem (1981) has instigated a new line of research which has been termed gender schema theory. Individuals are classed as being 'gender-schematic' or 'gender-aschematic' depending on whether they view and interpret the world using gender stereotypes or not. Gill (1992) points out, though, that this approach still relies on the Bem Sex Role Inventory for classifying individuals as gender-schematic or gender-aschematic, and so makes few advances on previous work.

Role Conflict

From the work investigating sex-role orientation developed research looking at 'role conflict'. It was hypothesised that female sports participants experience conflict in trying to fulfil the two contradictory roles of 'athlete' and 'female' - the athletic role involving aggression, strength and competitiveness, and the female role involving submission, sensitivity and passivity. For instance, Butt (1987) cites some famous female tennis players who she suggests have suffered emotionally through their sport involvement. However, no evidence is provided to support the suggestion that role conflict was the cause. Furthermore, Jackson and Marsh (1986) have analysed empirical studies and found very low levels of role conflict in female athletes. Hall (1981) has summarised much of the relevant research and has concluded that the majority of female sports players are in fact psychologically healthier than non-participants. While the notion of psychological health resulting from sport is not uncontroversial, there is no reason to suggest that females are less able to benefit from

sport than males. Hall claims that further research into role conflict is "pointless and even damaging", as it reifies a problem which does not exist. Similarly, Allison (1991) suggests that by 'clinging' to ideas of role conflict, researchers "reinforce and continually re-establish the very stereotypes that many wish to eradicate".

Motivational Orientation

The widespread criticisms of both the sex-differences approach and the gender-role orientation approach provoked researchers to move away from the "simple male-female and masculine-feminine dichotomies" (Gill, 1995). Sport is often portrayed as a classic 'achievement situation' (e.g. Wann, 1997) and researchers have studied achievement orientations and the motivations that people have for taking part in sport. It should be noted that while this is described as a unitary approach there are different perspectives. For example, some researchers approach 'achievement orientations' from an individual differences perspective (e.g. Weiss and Chaumeton, 1992) while others adopt a social-cognitive perspective and incorporate social factors into their analysis (e.g. Gould and Petlichkoff, 1988).

Individuals have been found to differ in their stated reasons for taking part in sport. Some people state intrinsic motives such as pleasure and personal mastery, while others may state extrinsic motives such as social approval and material rewards (Harter, 1978). Another distinction is a motivation to do well compared to other people ('ego-involved') versus the motivation to do well relative to the difficulty of the task ('task-involved') as described by Nicholls (1984).

One general finding is that males tend to be more ego-involved and see winning as more important than females do, whereas females, if anything, tend to view personal goals as being more important than males (Gill and Deeter, 1988). Findings, however, have been very mixed and Gill (1992) suggests that the work on gender differences in motivational orientations has not managed to indicate a single personality construct which can adequately explain patterns of sports participation.

Whatever the differences in achievement motivations between the sexes are, it is generally accepted that these differences are cultural in origin rather than biological (Murray and Matheson, 1992). That is, individuals learn a particular motivational orientation and it is therefore (theoretically at least) possible to change. Differences in the ways that children are brought up will influence the achievement orientations that males and females adopt. For instance, boys are expected to succeed at sport and are, therefore, rewarded for doing so. It is clear then that the rewards and reinforcement provided by other people are dependent on those people's attitudes and beliefs regarding what they expect of the sexes. Again it can be seen that gender stereotypes are crucial in that they will influence the type of motivational orientations that are instilled in children.

Self-efficacy

Another major focus of sport psychology in recent years has been on self-perceptions. As with motivational orientations, researchers have looked for ways in which individuals differ and then applied this to the problem of female under-representation in sport by looking for differences between the sexes (Clifton and Gill, 1994).

McAuley (1992) sees self-efficacy as the belief that one's actions will bring about a particular outcome - "a form of situation-specific self-confidence". According to Lirgg (1992), there are three models which could be used to explain why people are motivated to take part in sport and which all have a form of self-confidence as a central factor. These are Bandura's theory of self-efficacy (1977), Eccles and Harold's expectancy-value model (1991) and Harter's model of perceived competence (1978).

These models have been used to explain why women are under-represented in sport, as researchers have generally found that females have lower self-confidence than males (Lenney, 1977; Lirgg, 1991). The problem is that each of these models proposes different pathways for the effect of self-confidence. For instance, Bandura proposes that self-efficacy directly influences attributions (e.g. perceptions of control),

while Eccles and Harold propose the opposite, for these authors perceptions of control directly influence self-efficacy (performance expectations). Harter suggests that perceptions of control and self-efficacy (perceived competence) are simultaneously influenced by other factors.

It is not enough to know that there are differences in self-confidence between the sexes. If self-confidence is part of the reason behind the patterns of female sport participation then it is important to understand the causes and specific consequences of lower self-confidence (Lirgg, 1992). All three models mentioned above suggest that socialisation in particular is a primary cause of the lower levels of self-confidence amongst females in sport. However, Lirgg suggests that identifying this cause is inadequate in itself, it is also important to investigate the mechanisms behind the socialisation process and the ways in which the effects on self-perceptions are brought about.

The self-efficacy approach also suffers from the problems associated with early sex-differences explanations of sports participation patterns. In particular, differences in self-confidence could be the *result* of differential sports involvement rather than the cause. In fact, recent research seems to suggest that sex differences in self-confidence are highly dependent on the context and the particular task being studied. A study investigating cheerleading (perceived by both cheerleaders and non-cheerleaders as being a feminine activity) found that females were more confident in their own ability than males (Clifton and Gill, 1994). Any explanation of female participation patterns must therefore also take into account the sporting context.

Expectations of success and the levels of self-confidence which participants show will vary if the tasks are usually associated with either males or females. Whether the activity is associated more with males than with females comes from gender stereotypes, and beliefs about what males and females 'should' do. So, just as gender stereotypes were proposed to be important influences on motivational orientations, they are also critical influences on self-confidence.

Attributions

Along with self-efficacy, another important facet of self-perception, or what McAuley (1992) terms "self-referent thought", is the way in which individuals attribute explanations to outcomes of achievement situations. Weiner (1985) proposes that individuals will attribute a wide variety of explanations to outcomes, but what is important are the underlying "causal dimensions". The most commonly found dimensions are locus of causality, i.e. whether the cause of the outcome is believed to lie internally (within the self) or externally (situational), stability, i.e. whether the cause is seen to be stable over time or not; and controllability, i.e. whether the individual sees the cause to be controllable or not.

Researchers have found some sex differences in patterns of attributions following success and failure. In general, females have been found to be more likely than males to attribute success to external and unstable causes, and failure to internal and stable causes (Blucker and Hershberger, 1983). Deaux (1984) suggests that these differences are in fact due to differences in the initial expectations of success. When expectations of males and females are similar, the attributions will be similar. Gill (1992) points out that in sport, expectations are very rarely similar for males and females. Blucker and Hershberger (1983) highlight the fact that while in some domains different attribution patterns have been reliably demonstrated, the sport literature remains inconclusive. One problem has been the lack of carefully constructed studies to directly compare male and female athletes.

The way in which people attribute causality to other people's behaviour is also relevant. The influence of parents and peers may be based on stereotypic attitudes towards males and females. When explaining other people's behaviour, unexpected outcomes are attributed to unstable factors, whereas expected outcomes are attributed to stable factors (Fiske and Taylor, 1984). Thus, if an evaluator expects females to perform poorly (through holding a negative gender stereotype), the attributions used to explain a female's success may be unstable and outwith personal control.

Similar to the work on self-efficacy, the research on attributions can also be criticised on the question of causality and the subsequent loss of explanatory value, because the differences in males' and females' attributions regarding success and failure at sport could stem from the different experiences that males and females have of sport.

There is no doubt, though, that all of these areas of work: self-confidence, attributions and motivational orientations, are going to have some influence on individuals' sport behaviour. One of the problems facing sport psychology is to integrate these different approaches as it is unlikely that one of these factors alone will be able to account for sports participation figures (Weiss and Glenn, 1992; Plaisted, 1995). However, all these approaches hypothesise that sex differences on these measures are partly (if not wholly) caused by cultural and social influences. Socialisation research has investigated the role of social factors on individuals' sports involvement. As will be seen, socialisation research is similar to the previous approaches, in that gender stereotypes are an important feature of the explanatory model.

Socialisation

One definition of socialisation is "the process of social interaction through which people develop, extend, and change their ideas about who they are and how they relate to the world around them" (Coakley, 1987). Another definition is "the business of learning the normative standards of society" (Lee and Newby, 1983). Integral to sport socialisation research is the assumption that a person's choice of sports will be influenced by what they believe society expects of them.

Research which has investigated how people become involved in sport has found that, generally, teachers, peers and coaches are the main socialising agents (Lewko and Ewing, 1981; Lewko and Greendorfer, 1988; Greendorfer, 1979). Further research has identified the socialising agents which are most influential to males and those which are most influential to females (e.g. Greendorfer and Lewko, 1978) but has not attempted to explain the mechanisms by which these influences work. For instance,

while it is known that parental beliefs may lead to daughters being treated differently from sons, it is not known how this comes to influence children's self-perceptions (Brustad, 1992). While some psychological approaches have attempted to explain this, research so far has proved inconclusive (e.g. psychoanalytic theory, social learning theory) and the area remains controversial (Greendorfer, 1992a).

Socialisation research has found that from an early age, socialising agents such as parents, peers, siblings, teachers and coaches, all influence not only what sports people become involved in but also participants' perceptions of ability and enjoyment (Horn, 1987; Greendorfer, 1992b). However, socialising agents' influences are often based on stereotypic views of the sexes. A parent's influence might be vitally important in shaping the self-perceptions of a child, but if the parent holds strong views about males and females in sport then this influence may result in boys and girls having very different self-perceptions. If it is through socialisation that girls and boys develop different motivational orientations, levels of self-confidence and attributions then it is important to understand the mechanisms behind this influence. Part of understanding these mechanisms includes understanding exactly what people believe about males and females in sport.

Gender Stereotypes : The Common Link

Although there has been a variety of approaches which have attempted to explain participation patterns in sport, approaches which identify differences between males and females on particular measures (e.g. intrinsic or extrinsic motivation, high or low self-efficacy) lack explanatory value because the differences could arise as a result of different sports involvement. If female involvement in sport is to be increased then it is important to examine the processes which lead males and females to exhibit these differences and to establish the direction of causality. Gender stereotypes are important in these processes and it important to understand more about the relationship between the process of stereotyping and the processes which influence

sport participation. Before these relationships can be extricated, it is necessary to establish what people believe about males and females in sport.

As previously noted, each approach that attempts to explain sports participation patterns suffers from its own particular problems and together they suffer from a lack of integration (Landers, 1983; Plaisted, 1995). One factor which does link these approaches is the importance of gender stereotypes. Researchers looking at self-efficacy, attributions and motivational orientations all accept that social factors, under a variety of different names (e.g. socialising agents, society, cultural milieu), are exceptionally influential in creating the sex differences found on those measures. It is a central tenet of the present research that the influence of social factors, in all these different approaches, is often based on stereotypic views of the sexes.

While research into gender stereotypes and their acquisition has been extensive, empirical research into gender stereotypes in sport has been limited. In particular, studies investigating gender stereotypes in sport have been atheoretical and findings have been largely descriptive.

Studies by Weinberg, Reveles and Jackson (1984) and Parkhouse and Williams (1986) both used a disguised attitude measure (see Chapter Two) to assess attitudes towards male and female basketball coaches. Results from both studies show that male coaches were regarded most positively by male and female athletes, but female coaches were rated less positively, especially by male athletes. The main problem with these studies is that while the affective components of participants' attitudes were assessed, i.e. how positive or negative they felt towards the hypothetical coaches, their beliefs about male and female coaches (the cognitive components) were not assessed. It would seem reasonable to suggest that research into attitudes towards males and females in sport should not neglect any components if the process of stereotyping is to be fully understood. There is also the problem that only the attitudes of basketball players were assessed. There are many reasons to suggest that the attitudes of serious sport participants towards aspects of their sport are likely to be quite different from the attitudes of non-participants.

Wolfson, Ball and James (1985) also utilised a disguised attitude measure to assess university students' attitudes towards male and female sports participants who were described as jogging for different motives - either to improve their appearance or to improve their competence. Results showed that the male competence-motivated runner and the female appearance-motivated runner were rated most positively. The most negative attitudes were shown towards the male runner who was described as running to improve his appearance. The authors conclude that males who do not conform to gender stereotypes may be 'punished' more severely (by negative evaluation) than females who violate expectations. While this study does illustrate that males as well as females are subject to stereotypic expectations, one problem is the sample used. It is unlikely that the attitudes of university students (who were all studying sport studies) can be generalised to the wider population. In order to gain a greater understanding of participation in sport it is important to investigate the attitudes of non-participants as well as participants.

Bird and Williams (1980) studied the attributions for sport performance made by over 300 school children aged between 7 and 18 years. This sample represents a wide range of academic abilities and socio-economic backgrounds (compared to college student samples) and so this research is likely to be more informative regarding general societal attitudes towards the sexes in sport. Their results demonstrated that by the age of 13 all children tended to attribute male success in sport to effort, and by age 16 children tended to attribute female success in sport to luck. Bird and Williams found no evidence to suggest that males and females held different views of the sexes. This indicates that perhaps gender stereotypes reflect societal values rather than just one segment of society holding stereotypic views. While the sample used in this study was probably more representative of the larger population, the problem still remains that only a small part of gender stereotypes, namely attributions for success and failure, was investigated. While beliefs about why males and females are successful are no doubt an important component of gender stereotypes, these beliefs nevertheless comprise just one element.

Another area of research that has investigated gender stereotypes in sport (although indirectly) is that considering the sex-typing of sports. The original work in this field was carried out by Metheney (1965) who proposed criteria to determine the appropriateness of physical activities for females. A sport is described as 'sex-typed' if it is considered more appropriate for one sex than the other. Metheney suggested that sports involving body contact (e.g. wrestling) or transference of energy to a large object (e.g. javelin, discus) were seen as inappropriate for females, while sports involving aesthetic movement (e.g. gymnastics) were seen as appropriate. Matteo (1986) asked participants to rate the appropriateness of 68 sports and produced a 'league table' of sports with boxing seen as the most 'masculine' of sports, and cheerleading seen as the most 'feminine'.

Within this body of work, the relationship between stereotypes of the sexes, and stereotypes of sports is unclear. Gender stereotypes contain information regarding typical behaviours, while sport stereotypes may contain information regarding typical participants. There may be considerable overlap in the type of information the different stereotypes contain. Even if this problem remains unresolved, data on sex-typing of sports remains descriptive. What is important is the way in which gender stereotypes and sports stereotypes are used to influence sport behaviour.

There are various problem with current knowledge of gender stereotypes in sport. It is known that gender stereotypes are highly variable across cultures (and within cultures) and yet North American findings have been generalised to other contexts. There is good reason to doubt the relevance to other countries of knowledge about gender stereotypes in the U.S.A., given the unique legislation (Title IX) regarding females and sport, which has had enormous impact since implementation in 1972. Title IX demanded that any educational programme receiving federal funds must not show preferential treatment on the basis of sex. The result of this legislation was a massive increase in the amount of resources available for female sports (Hargreaves, 1994). Furthermore, the widespread use of American college students as participants in research looking at gender stereotyping restricts further the generalisability of results.

It is known that even very young children have strong beliefs about gender and gender-appropriate behaviour (Gill, 1995), and that children's experience of sport has a major bearing on sports involvement as an adult (Greendorfer, 1992b). Also, socialisation research has identified the peer group as a major influence on children's behaviour, and yet very little is known of children's attitudes towards the sexes in sport. Another problem with our knowledge of gender stereotypes in sport is that despite it being known that stereotypes change over time, results from the 1970s have not been updated to take into account the cultural and social changes of the past 20 years.

However, one of the most significant problems with past research into gender stereotypes in sport is the lack of clear theoretical frameworks. As will be demonstrated in the next section, mainstream psychology has provided many different definitions of stereotypes and the process of stereotyping had been conceptualised in radically different ways. Previous sports research has not distinguished the different perspectives on stereotypes and stereotyping and this limits the usefulness of findings. Sport psychology research must adopt appropriate theoretical frameworks if the impact of gender stereotypes on sports participation is to be fully understood.

Summary

Theories attempting to explain sport participation figures have neglected the role of gender stereotypes, particularly amongst children. It is accepted that stereotypes vary over time, between cultures, and across different age groups, and that children's peer groups are significant socialising agents. However little empirical work has investigated children's attitudes towards males and females in sport, and the stereotypes of the sexes that they hold. It is not enough to know that stereotypes are important. It is essential to know what stereotypes children hold, and how these stereotypes influence sport behaviour. This is the focus of the present research. The next section will describe previous research into stereotyping and highlight the controversies which concern the field today.

Stereotypes and Stereotyping

Stereotypes have been identified as being among the causes of prejudice and discrimination against certain groups within society. Gender stereotypes in particular are considered to be amongst the barriers restricting choices and opportunities for females in employment and education as well as in leisure and sporting contexts. However, the processes by which stereotypes exert their influence are still unclear, and a variety of theories propose different models of stereotypes and stereotyping. As already demonstrated, it is important to discover more about the processes of stereotyping if the aim of equal opportunities in sport is to be achieved.

Definitions

To begin this discussion it is necessary to clarify what is meant by 'stereotype' as there have been many definitions put forward (Miller, 1982). The word 'stereotype' literally means 'rigid trace' and was originally used in the printing trade to mean a solid metal plate cast from a mould. The first psychological use of the term was by psychiatrists who used 'stereotypy' to describe the repetitive movements of schizophrenic patients. Walter Lippmann (1922) was the first to use 'stereotype' to describe a mental representation and coin the usage which is most familiar today in social psychology. He saw stereotypes as 'pictures in the head' which helped to make sense of a world too complex to represent in an objective and accurate way.

A more recent and specific definition is provided by Oakes, Haslam and Turner (1994) who describe a stereotype as "the collection of attributes believed to define or characterise the members of a social group". This however contrasts with the definition suggested by Fiske and Taylor (1991) who see stereotypes as "...a particular kind of role schema that organises people's expectations about other people who fall into certain social categories". A role schema is the cognitive structure that organises one's knowledge about the behaviour expected of a person in a particular social

position. Oakes et al. emphasise attributes i.e. personal characteristics or traits, whereas Fiske and Taylor emphasise behaviours expected of particular people.

Leyens, Yzerbyt and Schadron (1994) combine both of these points of view and describe stereotypes as "shared beliefs about person attributes, usually personality traits, but often also behaviours, of a group of people". The inclusion of the word 'shared' seems contrary to the view that stereotyping is an individual process. However, Leyens et al. suggest that this is a confusion between the process (which is individual) and the end-product (which is embedded in a social context). Furthermore, because it is not possible to tell when an individual stereotype becomes a socially shared stereotype, they prefer to maintain 'shared' as a part of the definition. Jones (1982) avoids this issue and states simply that a stereotype "is nothing more than a set of interrelated characteristics we impute to a given group and its members" without indicating whether it is 'we as individuals' or 'we as a group' who impute characteristics.

While Leyens et al. include expected behaviours in their definition, Ashmore, Del Boca and Wohlers (1986) exclude this component from their definition of gender stereotypes. They define gender stereotypes as "the structured sets of beliefs about the personal attributes of women and men" and they suggest that some gender-related behaviour may not be due to stereotypes as such, but may be due to other gender knowledge such as scripts. Scripts are defined by Ashmore et al. as "overlearned sequences of behaviours". The debate about whether certain behaviours are driven by stereotypes or not may be a question of definition rather than a fundamental difference in the hypothesised structure of gender knowledge.

A common feature of the various definitions of stereotypes is the view that a stereotype is a kind of 'schema' or cognitive structure. For instance, Hamilton and Trier (1986) define a stereotype as "a cognitive structure that contains the perceiver's knowledge, beliefs, and expectancies about some human group". This is consistent with Fiske and Taylor's (1991) definition of a schema as "a cognitive

structure that represents knowledge about a concept or type of stimulus, including its attributes and the relations among those attributes".

The present research adopted a definition similar to that proposed by Golombok and Fivush (1994) who see stereotypes as "organised beliefs about characteristics of all members of a particular group". In the present research 'stereotype' was defined as 'organised beliefs about the characteristics of members of a particular group'. This definition was chosen because it does not specify that these beliefs *must* be shared (although they can be) and so avoids the somewhat extraneous debate regarding 'social' versus 'individual' stereotypes. Also, this definition allows for 'characteristics' to be interpreted broadly. As previously mentioned, the prevalent view of gender and sport behaviour suggests that explanatory models must be complex and multifaceted. Some research has proved insubstantial by concentrating on only one aspect of gender characteristics such as personality traits. There seems to be little point in arbitrarily excluding certain gender beliefs from definitions of 'stereotype'. The nature of the present research was partly exploratory and, therefore, was concerned with *any* gender beliefs that may influence sport behaviour. Therefore, 'characteristics' was interpreted broadly to include information on physical appearance, occupations, behaviours, personality traits, causal attributions, attitudes and interests.

As well as avoiding the issue of shared versus individual stereotypes, this definition of stereotype does not specify that the characteristics must apply to *'all'* members of a group' as Golombok and Fivush suggested. Within a social group, it is possible to get subtyping of members so that a certain characteristic only applies to some individuals. The definition adopted by the present research was chosen so that a particular attribute could be taken to broadly characterise the members of a group but not necessarily every individual within that group.

Early Research on Stereotypes

The first published research to specifically investigate stereotypes was carried out by Katz and Braly (1933). Using a 'check-list' methodology they asked 100 college students to describe members of various ethnic groups including Jews, Turks, Blacks, English, Italians and Americans. Descriptions were elicited by asking participants to pick from a list of 84 adjectives those traits which were 'typical' of the particular group, and to choose as many which were required to adequately describe the group. Participants were able to choose as many or as few adjectives as they liked, and they repeated the process for each of the ethnic groups. Aside from the characterisations of each group, it is interesting to note the high degree of consensus amongst the raters. Katz and Braly suggested that this could not be due to participants having personal experience of all these ethnic groups, and so the demonstrated stereotypes must reflect public attitudes.

It is also interesting to note how easy it was for Katz and Braly to get participants to describe each ethnic group. There appeared to be no 'social desirability' effect (see Chapter Two) preventing people from describing blacks as "superstitious, lazy, happy-go-lucky, and ignorant". However, in contrast, Gilbert (1951) reported that some participants in a similar study refused to make generalised judgements of ethnic groups.

Katz and Braly's work influenced future research on stereotypes particularly because of the methodology which emphasised the content of stereotypes rather than the process of stereotyping. Oakes et al. (1994) point out though that while Katz and Braly have been blamed for the emphasis on content of stereotypes, this is somewhat unfair as Katz and Braly themselves in a 1935 paper stressed that it was the mechanisms underlying prejudice that were important.

Nevertheless, research maintained the use of the checklist methodology to investigate such things as the rigidity of stereotypes across time, and changes during periods of international conflict. One of the features of this research was that stereotypes were

equated with prejudice. Stereotypes were considered to be wrong and wildly inaccurate because prejudice was wrong. For example, Schoenfeld (1942) said that "to the extent that a stereotype corresponds to objective facts, it is not a stereotype at all".

This view was challenged by those who suggested that there may be a 'kernel of truth' in stereotypes, and that what stereotypes do is represent real differences between groups but in exaggerated form (Hoffman and Hurst, 1990). Research moved away from describing the content of stereotypes towards assessing the level of accuracy of the content. The main problem with this type of research is the lack of 'objective' methodologies for assessing accuracy. As Oakes et al. (1994) point out, any methodology is subject to the researcher's own values and this will influence the degree to which any stereotypes are found to be accurate or not. However, some researchers (e.g. Judd and Park, 1993) maintain that, by using a variety of different measures, it is possible to create criteria against which stereotypes can be compared to measure their accuracy. Oakes and Reynolds (1997) dismiss this suggestion and question whether *objective* criteria are possible, even in principle. Judd and Park (1993) suggest that methodological problems such as unbalanced sampling and the use of biased judges have prevented objective criteria from being established. In contrast, Oakes and Reynolds (1997) question the assumption that a 'true' assessment of a person comes from studying their individuality by means of personality tests and clinical judgements. They suggest that even with ideal methodologies, the measurement of stereotypes is context-dependent and relative to the situation. What appears to be a 'true' picture of a person will vary depending on where, when, how and by whom judgements are made. It is, therefore, impossible to assess the accuracy of a stereotype if so-called objective measures of people's characteristics vary so much.

As stereotype research moved into the 1970s, less emphasis was placed on stereotype content and stereotype accuracy, and investigators became more interested in the processes and mechanisms underlying stereotyping. Unlike previous research, the more recent work, such as social cognitive research, has made a distinction between

prejudice and stereotyping. Baron and Byrne (1994) define prejudice as an attitude (usually negative) towards members of a social group. Attitudes are commonly held to consist of at least two distinct components. Fiske and Taylor (1994) point out that while theorists differ in their definitions of attitudes, all models include both an affective, evaluative component and a cognitive component. If prejudice is defined as an attitude, a stereotype constitutes the cognitive component. This is consistent with the definition of stereotypes as schemas presented earlier.

Gender Stereotypes

A major strand of research has concentrated on describing the content of gender stereotypes as well as trying to explain how these stereotypes are learned. While early work on stereotypes focused on race and ethnic stereotyping, during the 1960s and 1970s, gender stereotypes became the subject of much research, mainly due to an increase in awareness of sex discrimination and its causes that the Feminist movement brought (Ruble and Ruble, 1982).

Content of Gender Stereotypes

Ashmore, Del Boca and Wohlers (1986) suggest that research assessing the content of gender stereotypes has utilised three main methodologies. The first of these is the open-ended description such as that used by Sherriffs and McKee (1957) who asked participants to list traits or characteristics of men and women. The second main methodology used adjective checklists, in which participants are presented with a list of adjectives and asked to indicate those which characterise men and those that characterise women (e.g. Williams and Best, 1975). Thirdly, rating scales have been used to assess the degree to which certain adjectives or traits are characteristic of males or females (e.g. Rosenkrantz, Vogel, Bee, Broverman and Broverman, 1968).

Overall, these three different methodologies have described the content of gender stereotypes in similar ways. In general, males are seen as dominant, independent, competitive and unemotional, while females are seen as affectionate, dependent, submissive, nurturant and sensitive (Ashmore et al., 1986).

Most of this research assessed the content of gender stereotypes so that the participants could then be asked to what degree they themselves conformed to stereotypic views of males and females (e.g. Bem's Sex-role Inventory, 1974). The framing of research into gender stereotypes from a personality perspective has led to some conceptual and methodological problems.

Some researchers have conceptualised gender-related traits as being at different ends of a continuum, so that the endorsement of 'male' traits necessarily precludes the endorsement of 'female' traits. This 'bipolar' model is in contrast to the 'dualistic' model which suggests that a person can possess masculine and feminine characteristics. Bem (1974) describes such a person as 'androgynous' - a concept which has been subject to much criticism (Hall, 1981). Ruble and Ruble (1982) suggest that the question of whether stereotypes reflect a bipolar or dualistic model remains untested and may depend on the specific characteristics being studied.

Another of the major problems with assessing the content of gender stereotypes in these studies has been the use of samples of college students to characterise the sexes. Ashmore et al. (1986) point out that if the claims that 'gender stereotypes are pervasive' are to be supported, samples from a much wider range which reflect larger populations must be studied.

A further problem with the personality-based research has been the assumption that stereotypes consist primarily of personality traits. As discussed previously, stereotypes also include information on expected behaviours, physical appearance, occupations and sexual orientation. Deaux and Lewis (1984) found that people rated as sex-typed in physical experience and role behaviours are assumed to be sex-typed in personality traits as well.

It must be remembered that other characteristics apart from sex, such as race and age, affect our perceptions of other people (Golombok and Fivush, 1994). The stereotypes do not just 'add together' in an algebraic fashion. Thus the stereotype of say a 'black male' is not the same as the stereotype of a male added to the stereotype of a black person. While in general there may be clear differences between stereotypes of the sexes it must be remembered that gender stereotypes do not operate in a vacuum, but they interact with many other personal characteristics and the social context. This is particularly important for self-categorisation theory and social judgeability theory, both of which highlight the importance of context in impression formation (see next section). However, while some research has attempted merely to document the content of gender stereotypes other research has concentrated on how these stereotypes are learned.

Acquisition Of Gender Stereotypes

It has been demonstrated that even very young children have strong beliefs about the sexes. Signorella, Bigler and Liben (1993) in a meta-analytical review found that children between the ages of 3 and 6 years old are more rigid than adults in their beliefs about the sexes. This rigidity decreases with age when other information about people overrides gender information. Martin (1989), in a study of 4 to 10 years olds, found that all the children used gender information to predict behaviour (toy choice), but it was only the older children who used counter-stereotypic information over and above gender information when making predictions. Children have been found to apply their beliefs about gender more to children their own age than they do to adults (Urberg, 1982). This suggests that friends play a crucial role in children's developing concepts of gender (Golombok and Fivush, 1994).

While studies of pre-adolescent children have shown a decrease in rigidity of stereotypes with age, during adolescence it is suggested that children adhere more strongly to beliefs about gender (Ruble and Ruble, 1982). Insecurities and self-consciousness related to physical appearance, combined with strong peer pressure,

may, in fact, cause children to endorse more traditional views of the sexes. It should be noted that adolescents' concern over their physical appearance may have particularly strong influence over choices of sports and physical activity.

There have been two main approaches which have attempted to explain how children acquire gender stereotypes. These are social learning theory (Bandura, 1977) and the cognitive developmental approach typified by work looking at gender schemas.

Social learning theory suggests that children learn about gender stereotypes and gender appropriate behaviour through two main mechanisms - reinforcement and modelling. Children learn what is expected of the sexes through positive reinforcement of appropriate behaviours and negative reinforcement or punishment of inappropriate behaviours. For example, a girl might be rewarded with praise and encouragement for playing with a doll, but be greeted with disapproval for playing with a tool-set. Differential reinforcement comes from parents, peers, teachers and other socialising agents, who have stereotypic expectations about the sexes. One finding which is consistent across many studies is that boys are discouraged more from sex inappropriate behaviour than girls are (Golombok and Fivush, 1994).

Children can also learn about gender-appropriate behaviour through modelling of others behaviour. That is, children will observe males and females, but only imitate the behaviour of same-sex models. Children will, however, generally only imitate those models which they believe to show behaviours appropriate to their sex (Perry and Bussey, 1979).

The cognitive developmental approach, however, suggests that it is structural cognitive changes which allow children to make sense of the world around them. Children's ability to organise and understand their world changes with age (Ruble and Ruble, 1982). Whereas social learning theory suggests that the acquisition of gender knowledge is a fairly passive process, the cognitive developmental approach emphasises the active processes that a child engages in to make sense of their world. Research within these two different approaches has concluded that modelling,

reinforcement and cognitive-developmental changes are all important processes in children's acquisition of gender knowledge, and Golombok and Fivush (1994) claim that "The gap between social learning and cognitive theories of gender development has narrowed to such an extent that it is no longer meaningful to separate the two".

General Theories of Stereotyping

The distinction between prejudice and stereotyping, and an increased focus on process rather than content, heralded more theoretical approaches to stereotyping. There are many different theories of stereotyping. The areas of research which have been most concerned with stereotyping are self-categorisation theory and models of impression formation. The former approach developed from social identity theory and is associated with inter-group relations, while impression formation models are usually associated with an information-processing perspective and tend to be more individualistic. These will be discussed later along with a newer third approach, social judgeability theory, which is proposed by Leyens, Yzerbyt and Schadron (1994). This theory appears to be promising as it attempts to blend the best parts of social identity theory and traditional impression formation models. First, some of the other theories of stereotyping will be discussed. While the psychodynamic theories and realistic conflict theory do not deal directly with stereotyping, researchers in these areas laid the groundwork for later theories of stereotyping.

Psychodynamic Theories

One of the main psychodynamic approaches to stereotyping was put forward by Adorno, Frenkel-Brunswik, Levinson and Sanford (1950) in their book 'The Authoritarian Personality'. Central to their theory are the psychoanalytic concepts of repression, catharsis and projection. Adorno et al. suggest that individuals raised in a

strict, repressive family environment suffer from an 'authoritarian syndrome'. They are unable to express their views and feelings and so frustration builds up. This frustration leads to aggression, and instead of this being directed at the source of frustration (usually the father), aggression is projected onto less-valued outgroups.

Adorno et al. used a variety of scales to measure anti-Semitism, fascism and prejudice. Authoritarian personalities were described as inflexible in thought and prone to overgeneralisations and stereotyping of social groups. Stereotyping was not central to Adorno et al.'s theory but the expression of stereotypes typified a particular kind of person.

This approach to stereotyping and prejudice has attracted criticism on both methodological and theoretical grounds. The scales used to measure prejudice were constructed so that a positive answer always meant being more prejudiced, and so confounded acquiescence with prejudice. Also, validity was tested by interviewing participants who had already completed the questionnaires. However, the interviewers knew what those participants' scores were and this could have influenced the outcome of the interviews (Leyens et al., 1994).

More importantly though, this approach does not account for socio-cultural influences on prejudice. It is difficult to use individual differences to account for the uniformity of prejudice within some cultures. This approach is also unable to predict who will be the target of prejudice and discrimination (Leyens et al., 1994). This approach concentrated very much on prejudice and did not empirically study stereotypes or the processes behind stereotyping. Instead, stereotypes were assumed to be held by those individuals who expressed prejudice, and little effort was made to disentangle the constructs of 'prejudice' and 'stereotype'. Nevertheless, unlike earlier work, psychodynamic theories did place stereotypes within a theoretical framework.

Realistic Conflict Theory

As with the psychodynamic approaches, stereotyping is not central to realistic conflict theory. However, this approach is important because realistic conflict was important in the development of social identity theory, which in turn gave rise to self-categorisation theory which does deal with the process of stereotyping directly.

The main principle of the realistic conflict approach is that prejudice arises from competition produced by a lack of material resources. This theory derives from the series of studies carried out by Sherif, Harvey, White, Hood and Sherif (1961) at summer camps for eleven-year-old boys. Sherif et al. created a conflictual situation by dividing the boys into two teams who each had their own base camp. In the initial phase, the boys took part in a variety of sports and activities and after the first week it was found that the groups were highly cohesive, despite the fact that the groups were initially generated by splitting up close friends. In the next phase of the study, competitive events were introduced with rewards for the winning team. Sherif et al. found that this quickly led to the boys holding strong negative stereotypes of the other group, as well as discrimination and direct acts of aggression against the other teams. In the final phase of the study, the boys were set tasks which required that the teams worked together towards mutually desirable goals. After six days, conflict between the groups had largely vanished.

Baron and Byrne (1994) point out that this research was carried out over a short period of time, with only male participants who were from very similar social backgrounds. Despite these limitations, Sherif et al. have demonstrated that competition can lead to conflict and to negative attitudes and stereotypes of other groups. This does not, however, explain stereotyping of groups where there is no competition over material resources, or indeed of groups which have had no contact.

Social Learning

Psychodynamic approaches and the realistic conflict approach are more concerned with prejudice rather than stereotypes and stereotyping. As discussed earlier, the social learning approach, typified by the work of Eagly and her colleagues, deals directly with stereotypes and their acquisition. It should be noted that while Eagly's work is primarily concerned with gender stereotypes, the social learning approach is a general theory that is also applicable to other forms of stereotyping. However, while social learning theory provides some useful insights into some of the mechanisms behind stereotype acquisition, this approach has gained less support as a general theory of stereotyping.

Eagly (1987) suggests that stereotypes of males and females reflect actual observed differences in males and females through their distribution in particular social roles. Previous work has found males to be perceived as dominant and wanting control, and females are perceived as altruistic and nurturant (Deaux and Lewis, 1984). Eagly denotes the traits ascribed to males as 'agentic' and those ascribed to females as 'communal'. Eagly suggests that certain roles in society reflect certain traits. For instance, being a 'homemaker' requires more communal traits. Because of the sexual division of labour, females are observed in the role of 'homemaker' more than males are, and thus females are observed to show communal traits more often than males. Conversely, males are more often observed in the role of 'business executive', and are thus observed demonstrating 'agentic' traits more often than females.

The social learning approach to stereotyping has been criticised on a number of grounds. Firstly, the differences which stereotypes suggest between males and females may be too small to observe directly. In a meta-analytic review of 'influenceability' Eagly (1987) found that there was a mean effect size of 0.3 which is considered to be too small for people to even notice (Hoffman and Hurst, 1990). Secondly there are measurable differences between the sexes which do not constitute part of the stereotypes. For instance, females have been found to make fewer speech errors than males, and the reported effect size is 0.7, and yet this does not constitute

part of either the male or female stereotypes (Hoffman and Hurst, 1990). Thirdly, this approach does not explain why there are stereotypes of males and females in general, rather than 'homemaker' and 'business executive' stereotypes (Hoffman and Hurst, 1990). Finally, this approach does not explain why the outgroup is nearly universally denigrated. If stereotypes are based on observation of real differences between groups then positive evaluations of the outgroup should be encountered as often as positive evaluations of the ingroup. However this is not the case (Leyens et al., 1994).

Hoffman and Hurst (1990) put forward the suggestion that stereotypes are not observations of true sex differences but are rationalisations generated to explain the observed sexual division of labour. This is similar to the suggestion of Leyens et al. (1994) that stereotypes are used in a manner consistent with an individual's theory about the world. This will be discussed in later sections.

Self-Categorisation Theory and Social Identity Theory

Self-categorisation theory (Turner, Oakes, Reicher and Wetherell, 1987) developed from and extended the social identity theory which was first proposed by Tajfel in 1969. Both theories are based on the suggestion that individuals define themselves in terms of group membership, and that this definition leads to distinct psychological effects on social behaviour (Turner, 1988). Social identity theory was mainly concerned with intergroup conflict, whereas self-categorisation theory is a more general theory of group processes.

As previously mentioned realistic conflict theory posited competition over *material* resources as the basis for intergroup discrimination and associated stereotyping of the outgroup. In an extension to this theory, social identity theory suggests that conflict arises from competition for *symbolic* resources. For instance, a motivation to increase one's self-esteem may be all that is needed to generate friction between groups.

There are three main processes which comprise social identity theory. These are 'social categorisation', 'social comparison' and 'social identity'.

Social categorisation is the process by which a perceiver structures their environment. In particular, dividing the social world into an ingroup and an outgroup provides a basis for self-categorisation. From experiments estimating the length of lines, Tajfel (1969) suggested that mere categorisation leads to accentuation of intergroup differences and intragroup similarities. Tajfel suggests that this is the basis of stereotyping. In a study by Tajfel and Wilkes (1963) participants were asked to estimate the length of 8 different lines, which all differed from each other by a constant ratio. One group of participants were presented with each line unlabelled. When the four shortest lines were presented to another group, they were labelled with the letter 'A', and the four longest lines with the letter 'B'. A final group were presented lines which were randomly labelled either 'A' or 'B'. Results showed that participants who were presented with lines labelled 'A' and 'B' in a predictable manner exaggerated the differences between the longest of the 'A' lines and the shortest of the 'B' lines. That is, intergroup differences were accentuated, even though all 8 lines increased in size constantly.

Tajfel maintains that this categorisation process is also partly responsible for stereotyping in social settings. Evidence to support this comes from Wilder (1978) who asked participants to listen to a tape recorded group discussion. After one person had spoken, participants were asked to guess what the reactions would be. Participants were told that the discussants belonged to the same group or to two different groups. Results showed that when the participants believed the discussants to be from two different groups they predicted the subsequent reactions would be more different. When the participants believed the discussants to be from one group, reactions were predicted as being more similar. Thus, mere categorisation can be seen to lead to stereotyping of social groups.

Tajfel also showed that categorising individuals into groups is enough to elicit intergroup bias. The 'Minimal Group Paradigm' experiments (Tajfel, Flament, Billig

and Bundy, 1971) generated the most basic of 'groups' by asking school children to rate a variety of unlabelled abstract paintings which are actually by Klee or Kandinsky. Supposedly based on their ratings, each participant was told privately that they were a member of the 'Klee Group' (participants therefore assumed that some people were told they were a member of the 'Kandinsky' group). After this categorisation, participants were asked to pick rewards from a matrix. This matrix consisted of pairs of numbers representing pennies - one number corresponded to the reward for an anonymous ingroup member, the other the reward for an anonymous outgroup member. The participants were able to pick pairs of numbers that resulted in fairness, ingroup favouritism or altruism. Later studies used matrices that allowed for fairness, maximum group difference or maximum joint profit. Participants were told that their awards would not affect what they themselves would receive. Results showed that participants demonstrated considerable bias in favour of the in-group. This was despite the fact that the participants were all members of the same class and knew little about abstract paintings. However, telling participants that they were in one group as opposed to another group was enough to generate discrimination against the outgroup.

These results were taken to show that once categorisation has occurred, social comparison results in the ingroup being evaluated more positively than the outgroup. Individuals are most likely to rate groups on dimensions that are particularly important to them. By identifying with groups in this way, and by evaluating these groups in a positive manner, individuals can increase their self-esteem.

For Tajfel (1981), stereotypes have five main functions. These are the 'cognitive' function of organising the world, the 'motivational' function of representing important values, the 'explanatory' function for explaining social events, the 'justification' function for justifying actions, and finally stereotypes maintain the positive evaluation of the ingroup. One of the criticisms of social identity theory has been the lack of empirical work designed to test these functions.

Self-categorisation theory developed from social identity theory, and emphasises the importance not just of social identity but also individual identity. Turner and Oakes (1989) suggest that these social and individual identities are interdependent and must be analysed together.

It is important to note that for both self-categorisation theory and social identity theory, categorising individuals into groups is not necessarily seen as a maladaptive or erroneous process. While fitting someone into a category may indicate a move away from objective reality, Tajfel (1981) claims that these judgements actually help individuals to deal with their social world, by bringing '...subjective order and predictability to what would otherwise have been a fairly chaotic environment'. Furthermore, if categorisation is an error there must be some criteria by which it can be judged as accurate or not. Turner (1988) suggests that there is no 'one-way' to perceive the world as individuals construct their own social reality. Thus, the consequences of categorisation, i.e. stereotypes, are not seen as necessarily inaccurate. Accuracy will depend completely on the social context. Oakes et al. (1994) provide the example of a demonstrator taking part in the 'poll-tax riots'. For that individual it would not be useful to perceive a policeman as an individual. It would be useful for that social interaction to stereotype (categorise) the policeman because in that situation the demonstrator would need to know that this person, a policeman, would be likely to carry a truncheon, have the power to arrest people and charge them. In this context it is perfectly valid to stereotype that person, because it is more useful to perceive him as a member of a group than it is to perceive him as an individual.

Self-categorisation theory suggests that how people categorise individuals depends on the social context. That is the person being categorised may remain completely unchanged but be categorised differently depending on the terms of reference. For example, psychologists may be categorised as 'cold, hard scientists' when compared to sociologists, but as 'wishy-washy pseudo-scientists' when compared to physicists.

Self-categorisation theory proposes that the categorisation that becomes salient depends on the full range of stimuli being considered and the principle of 'meta-contrast'. This principle states that "...a given set of items is more likely to be categorised as a single entity to the degree that differences within that set of items are less than the differences between that set and others within the comparative context" (Oakes et al., 1994). For example, an individual might be categorised as an 'Australian' if the differences between Australians are perceived to be less than the differences between 'Australians' and 'Americans'. Alternatively, that same person may be classed as an 'English-speaker' if the differences between English-speaking groups (e.g. Australians and Americans) are seen as less than the differences between English-speaking groups and non-English-speaking groups. Meta-contrast is not the only factor in determining the basis of categorisation. The category must also fit the properties of the stimulus, and the social meaning of the category will determine whether it is applied. Oakes et al. (1994) term this the 'normative fit' of a category.

The important principles of self-categorisation theory and of social identity theory in relation to stereotypes can be summarised as follows (Leyens et al., 1994) :

1. Stereotypes are products of a "cognitivo-perceptual process" which accentuates intragroup similarities.
2. Stereotypes explain and justify behaviours
3. Stereotypes define behaviours, attitudes and perceptions appropriate to group membership.

The main problems facing social identity theory and self-categorisation theory revolve around the concept and measurement of self-esteem. One of the central tenets of social identity theory is that individuals are motivated to increase their self-esteem. Studies which have tried to demonstrate higher self-esteem because of ingroup favouritism have failed to adequately distinguish between global self-esteem and specific self-esteem, or provide valid ways of measuring self-esteem (Hogg and Abrams, 1988).

A further problem for social identity theory is that results from the minimal group paradigm studies were used to explain discrimination but the role of interdependence on participants' allocations of rewards has not been adequately established (Rabbie, Schot and Visser, 1989). Participants may allocate rewards to those people they feel dependent on rather than those they identify with. Questions have also arisen about the relevance of this paradigm to everyday life, as prejudice and discrimination are quite different phenomena from allocating rewards to anonymous group members, and there is no evidence to suggest that the participants actually stereotyped the outgroup members.

Despite these problems, social identity theory and self-categorisation theory have advanced greatly the study of stereotyping. The most important contribution of these two related theories is that stereotypes are not considered to be errors. According to Oakes et al. (1994:p187) "...stereotyping is psychologically rational, valid and reasonable.....it reflects reality accurately". This view is in contrast to the other dominant approach to stereotyping - impression formation.

Impression Formation

Within 'social cognition', a perspective which emphasises the processing of social information, impression formation and person perception models have been the theories most concerned with stereotyping. These approaches are characterised by a focus on processes rather than on outcomes, and adopt paradigms and methodologies from mainstream cognitive psychology (Fiske and Taylor, 1991).

Like the social identity theorists, impression formation researchers were dissatisfied with previous work that had merely described the content of stereotypes. Instead models of impression formation were posited that concentrated on *when* stereotypes are used. The two main models of impression formation are the continuum model (Fiske and Neuberg, 1990) and the dual process model (Brewer, 1988).

The continuum model suggests that the processes of forming impressions lie along a continuum with 'categorisation' (i.e. stereotyping) at one end, and 'individuation' at the other. The process of forming an impression of someone starts when the perceiver categorises the initial information about that person. Unless the perceiver is motivated and allocates more cognitive resources, the process stops here and the person has been categorised or stereotyped. If the perceiver does, however, devote more cognitive resources to attending to this information the process moves down the continuum to 'confirmatory categorisation'. In this process the category is compared to the data about the person. If the 'fit' is good or the perceiver is not motivated to consider the information further then, again, the process stops. However, the process may continue down the continuum to 'recategorisation'. Once more the process may stop, or move down to the other end of the continuum where the characteristics of the person are integrated to form an individual impression of that person. Movement down the continuum depends on sufficient cognitive resources being devoted to recategorisation by a motivated perceiver.

Brewer (1988) proposes a similar model in that perceivers either base their impressions on categorisation or on individuation. However, rather than these processes lying on a continuum, Brewer proposes two distinct paths. Initial identification of a person occurs automatically and the process ends if the person is not relevant. If the person is relevant and the perceiver is motivated, the identification is based on the personal attributes of the target individual. If the perceiver is not sufficiently motivated, categorisation occurs using prototypes, or if there is a poor match between the incoming information and the category invoked, then individuation takes place.

There are several important differences between these two models, but both envisage a tension in impression formation between stereotyping the target on the one hand, and considering them based on their personal characteristics on the other. This is similar to social identity theory which says that people can be treated as members of a group or as individuals depending on the social context. The primary difference between these two approaches is that for impression formation models, relying on

stereotypes is considered to be inaccurate. Fiske and Taylor (1991:p136) typify this view :

The litany of schematic effects on inferences and evaluations makes one wonder if people are not attuned to reality, leaning on expectancies and schemas, rather than confronting the data.

The reason given for people relying to heavily on stereotypes is that as perceivers we lack motivation to devote enough cognitive resources to be able to perceive someone as an individual with all their characteristics. The 'cognitive miser' model suggests that perceivers have a limited information processing capacity. To be efficient shortcuts must be used to maximise available cognitive resources. Stereotypes and schemas are viewed as ways of reducing the amount of information being processed.

The 'cognitive miser' model of social perceivers ignores almost entirely the role of motivation in impression formation (Fiske and Taylor, 1991). As the importance of motivation has been demonstrated, the 'cognitive miser' model has 'matured', and for Fiske and Taylor (1991) the prevalent view now of the social perceiver is as a 'motivated tactician'. Rather than always choosing the quickest or most efficient cognitive 'shortcut' for perception (as the cognitive miser model would suggest), the 'motivated tactician' has a variety of cognitive strategies which can be used at different times depending on motives and needs. The following quote is illustrative of this more recent perspective:

Sometimes the motivated tactician chooses wisely, in the interests of adaptability and accuracy, and sometimes the motivated tactician chooses defensively, in the interests of speed or self-esteem. (Fiske and Taylor, 1991:p13)

It is clear that for the 'motivated tactician' model the more that people rely on stereotypes the more inaccurate they are being. A 'wise' decision is equated with 'accuracy'. This is directly opposed to the view of social identity theorists who suggest

that there is no one objective reality and, therefore, no criterion for accuracy. Oakes et al. (1994) suggest that while impression formation models see stereotyping as an unfortunate side-effect, social identity theories view stereotyping as the "fully intended outcome of the categorisation process". For social identity theory, group membership and all that involves is a very real part of an individual's identity. There are situations when it is perfectly valid to stereotype individuals and to treat them as members of a group. There are also times when it is invalid to stereotype individuals as this leads to prejudice and discrimination, but it is the social context that determines this. In prejudice, it is when the stereotyping process is used that is wrong. The cognitive miser model and the motivated tactician model suggest there is something inherently wrong with the process of stereotyping itself. Social judgeability theory (Leyens et al., 1992, 1994), a more recent development in stereotyping research, also opposes the view that stereotypes are necessarily inaccurate, and suggests that pragmatism should replace accuracy as the criterion for judging how 'good' social perceptions are.

Social Judgeability Theory

The Best of Both Worlds: Combining Social Identity Theory and Impression Formation Models

Leyens et al. (1994) propose social judgeability theory as the best way to look at stereotyping as it blends parts of the social identity/self-categorisation approach and the impression formation models. For these researchers the process of stereotyping is "the process of applying a - stereotypical - judgement" such as rendering individuals interchangeable with other members of their category (Leyens et al., 1994). For perceivers to make a judgement they not only have to be able to fit a category to the information they have about the person, but they must also feel that they are in a position to judge. For Leyens et al. the important question is when will perceivers feel able and sufficiently confident to make a judgement.

Four criteria ('levels of adequacy') are suggested as being important when considering judgements. The first of these is 'reality'. Impression formation researchers suggest that this is the only important criterion when making judgements of others. Leyens et al. accept that reality is important as, after all, if a judge and the person being judged "do not share the same delusion", normal social interaction would not be possible (Leyens et al., 1994). The problem for the perceiver in making a judgement is to integrate category information with individual information of a target so as to match reality as well as possible. In addition, however, the perceiver needs to be able to make a judgement that is socially useful. There is no point in making a judgement if it does not make functional sense, so the other criteria to be considered deal with perceivers' theories about their judgements.

The 'integrity' level of adequacy suggests that a perceiver makes a judgement of others so that the integrity of their self and of the groups they belong to remains intact. This is similar to the Social Identity view that individuals downgrade the outgroup to maintain their own self-esteem.

The 'cultural' level of adequacy suggests that people respect social rules when making judgements. These rules vary across cultures and across time. For instance, Leyens et al. suggest that in Western culture, at this time, there is a social rule which says it is wrong to make judgements of other people based purely on categorial information.

The 'theoretical' level of adequacy suggests that when people make a judgement of another person, this judgement will correspond to some theory that the perceiver has about the world.

Leyens et al. (1994) sum up their social judgeability theory by saying that

..an adequate impression of someone should not only match reality, respect certain social rules, and protect people's identities, it should also constitute an enlightening gestalt that gives meaning to the world and allows communication.

This is a pragmatic theory and unlike social cognition which views people purely as information-processors, social judgeability theory stresses the social context within which judgements take place. The authors of this approach see social judgeability theory as a 'middle-of-the-road' position and admit that all their levels of adequacy have been considered separately by previous researchers.

Leyens et al. (1994) have concentrated particularly on the cultural level of adequacy to demonstrate their theory of when stereotyping occurs. As previously mentioned, these authors suggest that in Western cultures, at the present time, there is a general social rule that says it is wrong to make a judgement about another person if all that is known about that person is which category they belong to. Stereotyping occurs when a perceiver feels confident that they are in a position to make a judgement, but when the information is only categorial (e.g. whether the target being perceived is male or female). When this happens any judgement can only be based on the target's category membership, and a judgement based on category membership is a stereotype.

Empirical Support for Social Judgeability Theory

Yzerbyt, Schadron, Leyens and Rocher (1994) suggest that one way of making participants feel more confident about making a judgement is to manipulate the 'meta-information' of the experimental situation. 'Meta-information' is defined as the "extra-content aspects of a statement". In other words, every statement presented to participants about a target individual contains some content. For example, the statement 'David is an engineer' informs the reader the name of the target and his profession. However, the meta-information is all the information that can be gleaned from the presentation of the statement that is not part of the statement's content. For instance, if no more information is presented about 'David' then the reader will know that information was conveyed but that it was limited. Meta-information refers to all the aspects of the situation and context in which information is presented and social perception takes place. Yzerbyt et al. suggest that changes in meta-information (but

not in actual information) can lead to changes in a participant's feeling of confidence and their entitlement to judge.

To test this proposition Yzerbyt et al. (1994) experimentally induced participants to feel that they had individuating information when in fact they did not, and then asked participants to rate a target. Their study consisted of three stages. In the first stage, participants were asked to listen to a audio tape recording of an interview between two people, one of whom (the interviewee) was the target individual (i.e. the person to be perceived). Participants were not told the exact nature of the study. Instead they were told that the study was looking at the processing of information in social settings. At the end of the tape recording, the interviewee mentioned their profession - this was the categorial information. The profession mentioned was varied (comedian, which was diagnostic of extroversion, or librarian, which was diagnostic of introversion) across two conditions.

The second stage of the study was a dichotic listening task in which participants heard two recordings, one played into each ear, and were asked to shadow (i.e. repeat out loud) the speech in one ear. This kind of task is difficult for novices and requires great concentration. Participants are usually only able to report that they heard someone speaking in the unshadowed ear, and may be able to report the language or the sex of the speaker. Generally participants are unable to report any of the content of the unshadowed speech, though studies have shown that unconscious processing of the speech does occur. A classic example of this is the 'cocktail party scenario' where an individual can be intently engaged in a conversation with one person, but be instantly aware of someone else mentioning their name at the other side of the room (Cherry, 1953). Leyens et al. used this phenomenon to manipulate participants' feelings of entitlement to make judgements about the target individual. At the end of the second stage of their study, they informed half of their participants that although they were not aware of the speech in the other ear they had in fact subconsciously processed the information, some of which contained extra data about the interviewee. In Yzerbyt et al.'s words, these participants were given the 'illusion of individuating information'. The other participants were not told anything about the information they had heard.

The third and final stage of the study involved participants rating the target individual on a variety of personality scales. It is important to remember that all the participants had received the same amount of information about the target, as they had listened to the same tape recording in stage one of the study. The information presented during the dichotic listening task was irrelevant. However, half of the participants were told that they had received more information about the target when in fact they had not. Results showed that these participants were much more confident in their assessment of the target, and tended to stereotype the interviewee more (comedians seen as more extroverted, librarians as introverted). Yzerbyt et al. argue that because these participants believed they had a lot of information about the target they felt confident enough to make judgements about him. However, because they did not really have any more information about the target other than his profession, these judgements were based on stereotypes of comedians or librarians, depending on the condition.

There is, however, the possibility that social desirability effects were responsible for these results. It may be that those participants that were told they were given more information felt that they should be in a position to judge because of what the experimenter said, rather than because they felt more confident in their entitlement to make judgements. To test this possibility, Yzerbyt et al. (1994) carried out a further experiment which replicated the first study except for two features. Firstly, there was no manipulation of the target's profession to reduce the number of experimental conditions. The target was described as a comedian in all conditions. Secondly, participants were divided into three conditions after the shadowing task. Two of these conditions were the same as in the first study. That is, one group went immediately onto rating the target, while the other group were informed that they had received extra information about the target individual during the shadowing task (i.e. they were given the illusion of individuating information). However, the third group were told that they had received information about the group 'comedians' in the unshadowed ear. Yzerbyt et al. suggested that if social desirability effects were responsible for the results of the first study, then informing participants that they had received extra information about comedians would also produce more confident answers and stereotyped ratings. On the other hand, if their theory was correct, informing

participants that they had extra information about comedians would not make participants feel any more confident in their entitlement to judge the target individual. Results demonstrated that participants were again more confident in their judgements and made more stereotypic ratings in the condition where they were informed that they had received information about the specific target individual. Importantly, participants who were informed that they had been given more information about comedians were not significantly more confident or stereotyped in their judgements than the participants who were not informed about receiving any information. These results, according to Yzerbyt et al., reject the possibility that social desirability effects could explain participants' increased confidence in making judgements, and lend support to the theory of social judgeability.

This theory has been used to reinterpret a study by Darley and Gross (1983). Participants were presented with a video of a girl who was depicted as either coming from a low socio-economic background or a high socio-economic background. Just showing the video was not enough to elicit different judgements of the girl. However, adding the same sequence depicting the girl taking an intelligence test (but not indicating how well she did) to the end of both videos did elicit biased evaluations. In particular, the girl was described as being more intelligent when she was portrayed as coming from a high socio-economic background than a low socio-economic background. Darley and Gross suggest that the initial category information (socio-economic background) was enough to activate stereotypes. However, the participants did not feel that stereotypes are a valid basis for making judgements. Instead, the stereotypes acted as hypotheses for the interpretation of the later information, and this resulted in the biased judgements. Alternatively, Leyens et al. suggest that, as in their own study, by presenting the extra information about the intelligence test, participants were given the illusion of individuating information and they therefore felt entitled to make judgements about the girl. However, because the participants did not really have any individuating information the judgements could only be based on stereotypes and hence the biased judgements of intelligence.

By blending different aspects of the major theories of stereotyping, social judgeability theory provides a useful framework for studying the processes of social perception. Self-categorisation theory is somewhat limited in its explanation of stereotyping by concentrating on the social or personal identities that people adopt, and neglecting the role of norms (Leyens et al., 1994). However, Leyens et al. (1994) suggest that there is no reason why social rules can not be incorporated into the social identity/self categorisation perspectives.

The key concept of social judgeability theory is that people will only make judgements about other people when they feel confident that they are in a position to do so, and the four levels of adequacy that Leyens et al. postulate describe clearly the criteria that need to be fulfilled. Thus, an important feature of this theory is that it can explain when stereotyping will occur and also provide reasons for situations when stereotyping does not occur.

The Present Research

Past research has identified psychological factors linked to sport behaviour and demonstrated sex differences on these factors. Research has also identified social factors which influence sport behaviour but there has been a lack of integration of approaches and psychological factors have been studied in isolation from other factors such as the influence of socialising agents. Children's experience of sport influences their involvement as adults, and one of the most influential socialising agents for children is their peer group. Furthermore, the influence of socialising agents is often based on stereotypic views of the sexes. Due to a lack of recent, U.K.-based research, very little is known about the stereotypes that children hold of males and females in sport. If more females are to become involved in sport, it is important to know more about the stereotypes of males and females in sport that children hold.

The primary aim of the present research was to investigate the *content* of the stereotypes that children hold of males and females in sport. As the research progressed, a further aim developed which was to investigate *when* children express stereotypes. As discussed earlier, social judgeability theory specifies the conditions under which stereotypes will be expressed, and so this further aim centred on the *process* of stereotyping. In particular, the predictions made by social judgeability theory, with regard to social norms and their constraint on children's confidence in making stereotypic judgements, was examined.

The five studies designed to carry out these aims are described in Chapters 3 to 7. Chapter 2 discusses the different methodologies used in these studies.

CHAPTER TWO: THE ASSESSMENT AND MEASUREMENT OF STEREOTYPES AND ATTITUDES

According to Dewar and Horn (1992:p21)

...most topics in sport psychology cannot be adequately addressed without using a variety of methods and interpretive strategies that will allow researchers to compare and contrast the results obtained through different ways of knowing.

In line with this view, the present research utilised two contrasting methodologies in an attempt to answer the overall research question regarding the stereotypes that children hold of males and females in sport. Studies One, Two, Three and Five employed an indirect measure of stereotypes - a questionnaire based on the paradigm developed by Goldberg (1968) for identifying the operation of gender stereotypes. The data from these studies were analysed quantitatively. Study Four, however, employed a semi-structured interview to assess children's attitudes towards males and females in sport, and the data from this study were analysed qualitatively.

Multiple Research Methods

There are strong theoretical reasons to suggest that combining different methodologies to answer a research question is not just desirable but essential (Brewer and Hunter, 1989). Some of these reasons are general and can be applied to many topics within the social sciences, and revolve around the concept of 'triangulation' to 'overcome the inherent weaknesses of single measurement instruments' (Denzin, 1978). Furthermore, there are also strong theoretical reasons why stereotyping research in particular should employ multiple research tools. Specifically, social identity theorists (e.g. Oakes, Haslam and Turner, 1994) and social judgeability theorists (e.g. Leyens et al., 1994) have highlighted the context-

dependency of the stereotyping process. By using different methods it is possible to investigate the process of stereotyping in different situations.

Multiple Methods in Social Science Research

Brewer and Hunter (1989) suggest that any single methodology is restricted in its power to fully answer a research question. For instance, experimental research is unable to identify causal influences conclusively when confounding variables cannot be eliminated from the experimental setting. Also, interviews conducted in the field are restricted by their small scale, and cannot be used to investigate large social systems. By using carefully selected multiple methods of investigation, Brewer and Hunter claim that it is possible to generate 'stronger evidence' to support a given theory, than would be possible using a single method. In other words, a theory that has been tested using different methods will have greater explanatory value than a theory tested using only one method. Gould (1988) goes so far as to say that researchers should "beware of those who employ one method or instrument" and that for the advancement of knowledge "diverse methods must be employed".

Historically, certain research problems have been tackled using very similar methods. Often researchers deliberately use the same method to test the replicability of original findings (Davis, 1995). While replications are useful, using the same method does not subject the theory to a rigorous test, as the same types of error will also be replicated. For example, Webb, Campbell, Schwartz and Sechrist (1981) suggest that social research has over-relied on surveys and this makes findings particularly vulnerable to errors caused by sampling problems and reactivity (i.e. problems associated with participants being aware of the topic being researched and 'reacting' to this knowledge). Webb et al. propose that while nonreactive measures are also prone to certain types of error, they should be used along side other methodologies to overcome the intrinsic shortcomings of a single methodology. The results of one methodology can be used to corroborate or contradict the results of the other. In

other words, the use of multiple methods reduces "inappropriate certainty" (Robson, 1993).

Multiple Methods in Stereotyping Research

As well as general concerns regarding the overuse of single methods in social science research, there are also specific reasons why stereotyping research in particular should utilise different but complementary methodologies. Recent research has concentrated on the process of stereotyping (Oakes et al., 1994; Leyens et al., 1994) rather than merely identifying the content of stereotypes as early research did (e.g. Katz and Braly, 1933). This has highlighted the context-dependency of the stereotyping process, and has changed considerably the early view of stereotypes as fixed, rigid and static. Instead, stereotypes are seen as being fluid and relative, and are expressed in different ways in different situations.

Diab (1963a, 1963b) demonstrated that small differences in the ways that stereotypes were measured could influence the specific content of stereotypes of certain groups. Using the Katz and Braly Checklist methodology (Katz and Braly, 1933), Diab manipulated the total number of groups being assessed so that participants assigned traits to either 12 groups or 6 groups (Diab, 1963a). The group 'Americans' was rated more positively when participants were rating 6 groups as opposed to 12 groups. Diab also found that 'Americans' were rated more positively when characterised immediately after (and presumably directly compared to) 'Russians' than when characterised after 'Germans'.

In a further study using the same technique, Diab found that the stereotype of 'the French' was generally negative and included traits such as 'selfish' and 'materialistic' when rated amongst 13 national and racial groups. When the stereotype for 'the French' was elicited along with only 4 other groups, which were commonly held to be unpopular, the stereotype was much more positive and included traits such as 'sociable' and 'artistic' (Diab, 1963b).

These studies demonstrate that stereotypes are not rigid or fixed. Instead, stereotypes represent relationships between different groups. If the comparison situation is changed then a group can be perceived in radically different ways. As described in Chapter One, the way in which changes in context affect social perception is the principle behind self-categorisation theory (Oakes et al., 1994). The meta-contrast principle takes into account *relative* perceived differences between individuals, and for Oakes et al. stereotyping is a "context-dependent process which serves to represent the changing nature of intergroup relations" (Oakes et al., 1994).

If the stereotyping process is so sensitive to the context in which it occurs, then multiple methodologies should be used to investigate the process. By using different methodologies to tackle the same research question, it is possible to compare and contrast the stereotypes elicited in different contexts. As illustrated by Diab's studies (Diab, 1963a, 1963b), even small features of the questionnaire can have a marked effect on participants' stereotypes. The use of multiple methodologies can prevent invalid conclusions being made about stereotypes.

Predictions from social judgeability theory (Leyens et al., 1994) also suggest that multiple methodologies should be used to investigate stereotyping. One of the central tenets of social judgeability theory is that people must feel entitled to make a judgement before they will express their perception of another person. Leyens et al. suggest that there are four criteria which need to be fulfilled for an individual to feel entitled. These criteria or levels are the 'theoretical level', 'integrity level', 'reality level', and the 'cultural level' (see Chapter One). These levels suggest that a 'good' judgement (i.e. one that an individual feels entitled to express) should correspond to an individual's general theory of the world, correspond to reality in some way, and maintain that individual's social and personal integrity. Also, individuals make judgements in line with cultural norms. Leyens et al. propose that one such norm in present day Western culture suggests that it is wrong to make judgements about other people if the only information known about them is their social category. For example, most people would be reluctant to make judgements about a person if they were only told that his name was 'David' as there is so little information available.

However, as described in Chapter One, Yzerbyt et al. (1994) have shown how it is possible to make people think that they have enough information to make a judgement, even though they have not been given what they would normally consider sufficient individuating information. It is by manipulating features of the experimental context that Yzerbyt et al. were able to elicit judgements (which were based on stereotypes) from their participants. In different situations or contexts, people may feel more or less entitled to make judgements about other individuals. By using different methodologies and eliciting stereotypes in different contexts, it is possible to study the factors that make people feel more or less entitled to make social judgements.

There are, however, other ways in which context can be varied to study the effects on stereotyping without using radically different methodologies. Yzerbyt et al. (1994) added a condition to their experiment to give participants the illusion of having been given extra information, and this was enough to elicit stereotypes. Diab (1963a, 1963b) changed the order in which groups were rated, and the number of groups that participants were asked to rate. Again, these fairly subtle changes in context were enough to elicit different stereotypes. Thus, it is possible to manipulate changes in context within a single methodology and find radical differences in the stereotypes expressed.

As well as varying the methodology employed, the present research utilised a more subtle manipulation. Study Five consisted of a questionnaire similar to that used in Studies One, Two and Three, but varied the way in which the information was presented in an attempt to vary the stereotypes elicited. This can be considered as a 'within methodology' manipulation of context, whereas the five studies together can be considered as a 'between methodology' manipulation of context.

Summary

Recent theoretical developments have highlighted the important effects of context on the stereotyping process. There are various ways in which context can be manipulated

to investigate these effects. Within a single methodology, small changes can be made to the way information is presented. The use of multiple methodologies creates radically different contexts for the study of stereotypes. In line with calls for the increased use of multiple methodologies in social science research, the present research utilised a questionnaire for Studies One, Two, Three and Five, the results of which were analysed quantitatively, and a semi-structured interview, analysed qualitatively, for Study Four. Within Study Five the effects of context on the stereotyping process were investigated by changing the way in which information was presented to the participants.

Combining Quantitative and Qualitative Methodologies.

Despite many researchers (e.g. Miles and Huberman, 1994; Dewar and Horn, 1994; Denzin, 1978; Brewer and Hunter, 1989) calling for the use of multiple methodologies in social science research, combining qualitative and quantitative data within a single research project is still a contentious issue. For instance, Smith and Heshusius (1986) suggest that the quantitative approach and the qualitative approach are incompatible, and they decry calls for co-operation between the two perspectives. Their argument is based on the different ways in which the two paradigms developed.

The quantitative paradigm developed from the methods of the natural sciences such as physics and chemistry. Researchers in these fields adopted a rational and positivistic approach to uncovering 'the truth' about the world, using experiments to test hypotheses about 'reality'. It was assumed that there was a single objective reality which could be measured reliably and predictably. To study this 'objective reality' it was considered imperative for an investigator to avoid influencing or biasing studies at all costs. These beliefs and assumptions about the way research should be conducted were adopted by psychology as well as by other social sciences, as the means by which human behaviour and experience could be investigated (Miller, 1962).

On the other hand, qualitative research developed in direct opposition to these assumptions. At the beginning of the 20th Century, the 'Chicago School' initiated anthropological research within a different paradigm (Denzin and Lincoln, 1994). These researchers suggested that human beings could not be studied in the same ways as atoms, gravity and the planets. Instead, the subject matters of the social sciences - behaviour, feelings, emotions, thoughts, values, attitudes - were subjective and ultimately could not be measured or assessed in an objective way. Qualitative researchers also suggested that it was impossible to separate the interrelationships between the investigator and the phenomena being researched (Smith and Heshusius, 1986). While quantitative research attempted to minimise the effects of the researcher to the point where any possible influence on the results could be eliminated, qualitative research developed methodologies which acknowledged the role of the researcher in the process of investigating social phenomena. To these ends, qualitative research employed methods such as unstructured interviews and participant observation, and was carried out 'in the field' rather than under controlled experimental or laboratory conditions (Hogg and Vaughan, 1995).

It is because of the development of qualitative methods in direct opposition to quantitative methods that some researchers claim that the two types of methods are incompatible and should not be combined within a research project. For instance, Smith and Heshusius (1986) highlight the difference in the meaning of the concept 'truth' to the two paradigms. Within the quantitative paradigm, there is the assumption that there is a 'real' social world which can be observed and described in an objective manner. Truth is assumed to be a direct correspondence between the words that a researcher uses to describe the world and the 'independent existing reality'. In contrast, within the qualitative paradigm, because everyone interprets and constructs the social world in different ways, if there is to be a single 'truth', then it can only ever be "a matter of socially and historically conditioned agreement" (Smith and Heshusius, 1986). In other words, everybody has their own view of the world, and each individual's view is no more or no less 'true' than any other view. It is because each paradigm has such different definitions of truth, that Smith and Heshusius deny that the approaches can ever be compatible. Because of views like

this, for many years qualitative research was portrayed as a direct alternative to quantitative methods, and each perspective was promoted to the exclusion of the other (Richardson, 1996).

It is important, though, to distinguish between attempts to combine different and contrasting methodologies and attempts to combine different paradigms. Researchers such as Smith and Heshusius (1986) do not actually condemn the linkage of qualitative and quantitative methods entirely. They suggest that researchers of a 'realist orientation' are perfectly able to utilise methods more usually associated with qualitative enquiry, and similarly, researchers of a 'naturalistic orientation' can resort to quantitative methods to supplement their investigations. However, this does not mean that the two paradigms are compatible or complementary. Research aims to find 'knowledge' and 'truth', but Smith and Heshusius point out the different meanings of these concepts for the two paradigms in social science inquiry. Roughly, quantitative researchers investigate social phenomena with the view that there is a single, objective truth, whereas qualitative researchers believe in multiple, subjective truths. According to Smith and Heshusius, one research project can utilise many different methodologies but ultimately a project can only be carried out from one perspective. It is possible to integrate different methodologies, but it is not possible to reconcile the different meanings of 'truth'.

Smith and Heshusius lament the demise of the conflict between qualitative and quantitative researchers. They suggest that too many researchers have combined different quantitative and qualitative methodologies and this has implicitly suggested that the different paradigms are compatible. For Smith and Heshusius, social scientists should centre their debates on the differences between the paradigms, as these issues are central to the purpose of social science research and the directions it should take.

However, Miles and Huberman (1994) suggest that the qualitative versus quantitative debate is an 'unproductive' argument. For these authors, co-operation between different perspectives is more fruitful than conflict, and they suggest a variety of ways

in which qualitative data and quantitative data can be linked usefully. For instance, qualitative data can help in the design of questionnaires or survey materials by identifying important topics and issues. Quantitative research can be used to test predictions generated by qualitative studies. Significantly, though, Miles and Huberman do not suggest that one type of methodology is more important than the other. Instead they point out that both types of methodology can be used for descriptive, exploratory and inductive purposes, as well as "explanatory, confirmatory and hypothesis-testing purposes" (Miles and Huberman, 1994).

As will be seen in later chapters, one of the main aims of the interview study in the present research was to test possible explanations for the results from the previous three studies, as well as attempting to generate new ideas about the language and vocabulary used by the children to express their attitudes and stereotypes. Thus, the present research adopts a view similar to that of Miles and Huberman (1994) and Brewer and Hunter (1989) in that while integrating two different kinds of data within a research project may still be contentious, weaknesses in single methods can be overcome by combining methodologies. This is especially true for the present research as it is becoming more apparent that stereotyping is context-dependent (Oakes et al., 1994; Leyens et al., 1994). By comparing the results from different methodologies it may be possible to learn more about the ways in which contextual factors influence the process of stereotyping. This perspective is not necessarily at odds with the views of Smith and Heshusius, but it can be said that, in their terms, the present research was carried out under a 'realist orientation'.

The Development of Stereotype and Attitude Measures

Historically, the measurement of stereotypes has been closely linked to attitude measurement. Original work did not distinguish between negative attitudes (or prejudice) and the stereotyping process. Nowadays, negative attitudes are not considered to be synonymous with stereotyping. For the purposes of the present research, a stereotype was considered as just one component of an attitude - the cognitive component (See Chapter 1). An important implication of this definition is that while an attitude entails a consistent evaluative response, a stereotype is a varied collection of beliefs and information about an 'object' (social or otherwise) which may produce a variety of evaluative responses depending on context (Greenwald and Banaji, 1995).

Much knowledge of stereotypes has come from research into attitudes or prejudice. That is, the content and nature of a stereotype are often implied from the attitudes that the stereotype is assumed to guide (c.f. Katz and Braly (1933) who assessed stereotypes and then implied attitudes). The measurement of stereotypes is inextricably linked to the measurement of attitudes, even though attitudes and stereotypes are different constructs. The present research, while concentrating on stereotypes and the process of stereotyping, also took into account the evaluative and emotional aspects of children's attitudes. It is not enough just to know what people believe; it is also important to know what they feel. Different people may believe the same facts about a particular group. However, while sharing common beliefs, those people might have very different feelings towards the group in question. It is important to take this into account as these feelings may be crucial to the processes of prejudice and discrimination.

The various ways in which attitudes and stereotypes have been measured can be categorised along two main dimensions. The first of these is the type of response that the participants are able to give, ranging from a completely free response providing data which is generally analysed qualitatively, to a completely restricted choice such

as selecting adjectives from a list, providing data which is generally analysed statistically.

The other main dimension is the level of 'obtrusiveness' or the degree to which individuals are aware of the stereotypes being investigated. As previously mentioned, the present research was based on two different methods - one indirect measure and one direct measure. The first of these was a disguised attitude measure, based on the 'Goldberg paradigm', which asked respondents to rate target individuals on Likert-type scales. Participants, therefore, had little freedom in their responses and were unaware of the attitudes being investigated. The direct methodology used was a semi-structured interview which meant that participants were aware of the nature of the investigation and had greater freedom in their choice of responses. The following sections of this chapter will discuss the different methodologies, both direct and indirect, used to measure attitudes and stereotypes, and will then explain the rationale behind the selection of the methodologies utilised in the present research.

Direct Methods of Stereotype and Attitude Assessment

The Checklist Methodology

As discussed in Chapter 1, the original research into stereotypes carried out by Katz and Braly (1933) asked students to describe various groups of people such as Jews, Blacks and Chinese, by ticking from a list those trait adjectives that they felt characterised that particular group. The adjectives selected most often were taken to define the content of the stereotype. Hamilton, Stroessner and Driscoll (1994) suggest that while this method became the most popular for identifying stereotypes it is subject to several shortcomings. The reasons for its popularity are its ease of administration, the simplicity of the task, and the speed at which several stereotypes can be assessed at one time. Furthermore, this method is easy to replicate and thus possible to compare stereotypes across time.

However, this method is restricted to assessing only the trait aspect of stereotypes and ignores behaviours and other aspects of stereotypes. In other words, this method only investigates the *content* of stereotypes rather than the *process* of stereotyping. Furthermore, the type of response that participants are able to make is highly restrictive. They are able to either include a trait or exclude it from the list. Participants are unable to endorse one trait to a greater degree than another - it is an all or nothing response. This means that certain traits that are only slightly characteristic of the group may be included in the stereotype. Also, by restricting participants to a list of traits generated by the experimenter (albeit with the help of panels of judges), other traits that participants feel characterise a group may be omitted from the final results.

The Problem of Reactivity

A direct measure of stereotypes like this faces the problem that because the participants are aware of the attitudes being investigated, the way in which they respond may be affected. This problem is termed 'reactivity', i.e. participants react to the knowledge that their attitudes are being studied. The way in which participants react will always be impossible to predict. Participants may know that racism or sexism is wrong and they may try to appear unprejudiced by giving answers that are 'socially desirable'. Participants may also try to 'please' the experimenter by giving the sort of answers that they think the experimenter wants. Of course, participants may often misinterpret the experimenter's questions and the purpose of the study. These problems are especially troublesome when researching sensitive areas such as stereotyping of males and females (Dyer, 1995). There is now much more awareness of sexism, and stereotyping is seen as a 'bad thing' (Brannon, 1996). Participants may try to appear egalitarian when answering questions which ask directly their attitudes towards the sexes. This may lead to an underestimation of strong attitudes towards the sexes. On the other hand, people may report very strong attitudes and stereotypes, not because they themselves agree with them, but because they feel that males and females are viewed by society in general in such a way. This would lead to an overestimation of strong attitudes towards the sexes.

Various techniques have been developed within attitude measurement to reduce social desirability effects. One major development in the measurement of attitudes was to ask participants not just whether they endorsed a particular statement or item, but the extent to which it was or was not endorsed. This development was initiated primarily to elicit more information about the strength of people's attitudes rather than as a solution to reactivity. However, by developing rating scales it was possible to include amongst other things buffer items and negatively phrased items which were designed to reduce reactivity (Dyer, 1995).

Rating Scales

Semantic Differential Scales

The Semantic Differential Scale was developed by Osgood, Suci and Tannenbaum (1957). Unlike the Likert scale which will be discussed later, each item is rated on up to twelve dimensions rather than just one. The dimensions are bipolar and each pole is labelled with adjectives reflecting the extremes of a continuum, while the positions between the poles are unlabelled. Typically there are seven positions in all for the participant to endorse on each dimension, including a mid-point which participants can use to indicate neutrality.

An example of the semantic differential scale is to ask participants to rate 'American Communist' (the attitude object) on the dimensions 'active-passive', 'hard-soft' and 'strong-weak'. Osgood et al. argued that a person's attitude toward an object is equivalent to the object's evaluative meaning, and so the semantic differential scale is usually introduced to participants as a method for determining the meaning of words. Triandis (1971) suggests that this may disguise the intention of the experimenter to a small extent and therefore reduce social desirability effects. Certainly, because the scales are more general and broad than those typically used for Likert scales, and because a participant's attitudes are assumed from their evaluations, this method is somewhat less direct than other methods.

This kind of scale was considered inappropriate for the present research primarily because it proved difficult to generate items with suitably labelled end-points. It was felt that children as young as 12 would not understand what it means to describe someone as 'active' or 'passive'. Although alternative words could be used, some have obvious 'opposites', while other words which were of importance to this investigation of stereotypes do not. For instance, 'big' and 'small' might be used to label the end-points of a semantic differential scale and it is unlikely that this would cause too many comprehension problems for children. However, it is much harder to generate antonyms for words such as 'aggressive', 'butch' or 'athletic'.

Equal Appearing Intervals

The equal appearing intervals method was developed by Thurstone (1928) and requires participants whose attitudes are being assessed to select from a list of about fifteen to twenty items those statements with which they agree. The list of statements is generated by giving a large number of statements to a group of judges who are asked to place the statements on an eleven point scale depending on whether they see the statement as being unfavourable or favourable towards the attitude object. The judges are asked to treat the eleven points on the scale as if they are equal intervals. This procedure does not ask the judges whether they agree with the statements or not; the aim is solely to establish to what degree each statement is in favour of the attitude object. The distribution of each statement is analysed and only statements rated with a high degree of consistency are retained. Statements which have a bimodal distribution are assumed to be ambiguous and are rejected, as are statements which are dispersed widely across the scale by different judges. The mean or median score is calculated for the remaining items and this is known as the item's 'scale value' (Fishbein and Ajzen, 1975). Thus, when the list of statements is presented to participants it is possible to determine an individual's attitude score by calculating the mean scale value across the items that they have endorsed. Unlike a Likert scale, participants only accept or reject statements, and so on the face of it, the Thurstone scale is an easier rating scale for participants to complete. On the other hand, this limited choice of possible decisions does not give participants the opportunity to rate

the degree to which they endorse a particular statement. However, the main reason that this methodology was deemed not suitable was because the main aim of the present research was exploratory, and a Thurstone scale, as well as being time-consuming to construct, produces only one overall attitude score.

Another problem with this method is the assumption that the judges view the favourableness of the statements in the same way as the participants. Fishbein and Ajzen (1975) suggest that under most conditions this assumption is valid. Triandis (1971), however, suggests that the assumption is only valid when the judges are not extreme on the particular attitude being investigated. If the judges are extreme, or if they sort the items on the basis of whether they personally agree with them or not (rather than just rating the favourableness of the statement), then the scale values can be significantly distorted.

An additional problem with the equal intervals method of assessing attitudes is that it is much harder than with other methods to disguise the attitude being measured, and thus reactivity is still a major problem. While Likert-type scales and semantic differential scales can be designed to be slightly less transparent, all these methods suffer to some extent from the problems of social desirability.

Likert Scale

Likert scales (Likert, 1932) or scales adapted from Likert's designs ('Likert-type' scales), are amongst the most reliable and frequently used rating scales in psychology (Hayes, 1993). Typically, participants are asked to rate a series of items relating to the attitude under investigation (the attitude object) on a five point scale with labels such as *strongly agree*, *agree*, *undecided*, *disagree* and *strongly disagree*. Once participants have completed the scales their responses for individual items are summed to give an overall attitude score, usually by assigning a score of 5 to *agree strongly*, a score of 4 to *agree*, a score of 3 to *undecided* and so on (Fishbein and Ajzen, 1975). So in a measure consisting of 20 items, scores for attitudes will range from 20 (least favourable) to 100 (most favourable).

One potential problem with this method of attitude measurement is that if a participant has a general tendency to agree with statements (e.g. because of social desirability effects) then their favourableness towards the attitude object will be overestimated. To combat this problem, Likert scales usually include negatively phrased items, which require the participant to endorse either *disagree* or *strongly disagree* to show their agreement towards the attitude object being investigated (Triandis, 1971). For example, to show a positive attitude towards sport in schools, a participant would need to either disagree or strongly disagree with the item 'sport should be kept out of schools'. Thus the attitudes of participants who tend to agree with all items will not be overestimated. Instead, their overall score will indicate neutrality. While this may not be a reflection of their 'true' attitude, it does prevent an inaccurate assessment being accepted as an exceptionally positive attitude.

The original method for designing attitude scales proposed by Likert (1932) included item analysis to remove items which correlated poorly with the overall score for the measure. This process was designed to increase the validity of the complete measure by only including items which contributed significantly to the total score. However, another possible method to combat social desirability effects in the use of Likert-type scales is to actually include 'buffer' items or items which are not used to compute the overall attitude score. Buffer items are usually less extreme statements about the attitude object being studied, or statements unrelated to the attitude object. When included in scales which are assessing attitudes towards sensitive issues, the exact nature of the study can be masked. If participants are less able to guess exactly what the researcher is attempting to measure, socially desirable answers may be reduced.

One problem with the use of Likert scales is that participants are asked to rate a long list of items without understanding in advance the full range of items that are going to be presented. A participant may use the most extreme category 'strongly agree' to rate one item, but then find later an item with which they agree even more strongly. This item is also rated in the 'strongly agree' category even though the participant's 'true' attitude is that one item is viewed quite differently from the other. Foddy (1993) suggests that this problem could be ameliorated by allowing participants a preview of

all the items before being asked to rate them, or by explicitly presenting the most extreme positive and negative items at the top of the list. This latter solution, though, entails a prejudgement as to which items are the most extreme and clearly this may vary for different people.

True Likert scales are comparatively rare in contemporary psychological research. However, scales based on Likert's original system (Likert-type scales) are more common. There are various problems with the original technique that made it an unsuitable scale for use in the present research. Primarily, a Likert scale is used to gauge the overall positive or negative attitude towards a particular attitude object. The aim of the present research was exploratory and an overall attitude score was not calculated.

However, a Likert-type scale was employed in combination with a disguised attitude measure to exploit the advantages of Likert's methodology. The primary feature of the present research that constrained the type of methodologies that were employed was the use of children as participants. It was important that any rating scale employed was straightforward to complete and able to be understood by children across the secondary school age range. Access to the participants was constrained by the school timetable and so a questionnaire that was relatively quick to complete was a further important requirement. For these reasons, it was decided to adopt the five point response scales developed by Likert. Likert's original scales ranged from *strongly agree* to *strongly disagree*. As described in the next chapter, the participants in the present research rated the extent to which they felt that words were *very good descriptions* or *very poor descriptions* of target individuals. Other questions were answered using similar five point scales with labelled end-points. The main advantage of using such scales was that the participants could respond quickly using a format that, according to the school teachers, most of them would have been familiar with already. Furthermore, by using such a scale, it was possible to analyse each item individually without producing an overall attitude score. This was important because the present research was, to a large extent, attempting to find which components of

gender stereotypes were the most important. Unlike a true Likert scale, the items in the questionnaires in the present research were assumed to be independent.

Free-response Measures

A problem with all of the different rating scales is, that while they are able to assess the evaluative component of attitudes, they are less successful at determining the cognitive component of attitudes, or the content of stereotypes. Participants are only able to endorse, to a greater or lesser degree, items or statements which have been preselected by experimenters, albeit with the help of judges in some instances. Free-response measures are required to establish new ideas about the content of stereotypes.

To generate new ideas about the content of stereotypes a less restricted response format is required. Despite this, only a small number of studies investigating stereotyping have allowed participants a 'free response' (Oakes et al., 1994). Of these studies, the open-ended approach has generally been used in the first stage of constructing a standardised measure of self-stereotyping such as the Bem Sex-Role Inventory (see Chapter One), which gauges the degree to which participants feel they themselves conform to traditional stereotypes of males and females (Ashmore et al., 1986). Bem asked 100 college students to select from a long list of adjectives those traits which they believed were desirable for one sex more than the other. The original list was generated by asking participants to think of and list as many traits as they could.

Sherriffs and McKee (1957) asked participants to list 10 traits or characteristics of men and 10 traits or characteristics of women. The data collected was content analysed and produced 26 categories of similar traits and characteristics. For males, the most frequently mentioned traits or characteristics were characterised as 'physical attributes', 'ascendant' and 'responsible'. For females, the most frequent descriptors were categorised as 'social awareness', 'negative affect or behaviour towards others'

and 'physical attributes'. One of the problems with this kind of research is the lack of any one standard method of analysis. Comparisons across cultures and across time are very difficult to make, and this has contributed towards the popularity of quantitative methodologies.

Despite these problems, a methodology utilising a free response format still provides the most potential for generating new ideas which the experimenter may not have considered. For this reason, a free-response measure was included in the pilot questionnaires for Study One. The children were asked to write down anything they could think of that had not been covered in the questionnaire. However, as discussed in Chapter Three, this caused problems for some of the children who were unable to come up with any new ideas, and also excessively extended the time required for the children to complete the questionnaire. Because of these problems, the free-response measure was not included in the main questionnaire studies. However, Study Four did permit the participants to give their own answers to the questions and to express their views in their own words.

Interviews

Apart from the item generation and selection procedures used in the preparation of rating scales, interviews can also provide a free-response format within a direct assessment of attitudes.

According to Robson (1993) an "interview is a kind of conversation...with a purpose". Denzin also sees an interview like a conversation, though his definition of an interview - "an informal interchange of thoughts by spoken words" (Denzin, 1978) - ignores the fact that some interviews are highly formalised and the transmission of views may be unidirectional (e.g. structured opinion-poll interviews). However interviews are defined, they can be characterised by the collection of verbal data, in response to spoken questions. While this may seem a very broad definition, it does encompass the many different types of interviews, which range from highly structured to completely unstructured 'conversations'. Interviews can occur between one

interviewer and one or several respondents, in a face-to-face situation or over long distances by telephone or other media. Just as there are different ways of conducting interviews, there are different ways to record the verbal data. For example, an interviewer may take notes during an interview, write notes after the interview has finished, use a tape-recorder and then transcribe the tapes in part or in full later, or indeed, a combination of any of these methods.

Structured Interviews

A highly structured interview is akin to a verbal questionnaire in that the interviewer has a set number of questions, and these are asked in a predefined order - the interview 'schedule' (Robson, 1993). Often, great emphasis is placed on the interviewer 'controlling' the interview by asking the questions in the same way for all respondents and by refusing to reinforce or discourage particular responses so as not to lead the respondent. Importantly, respondents have little or no freedom in the type of response that they can give. Often, the interviewer presents a series of options from which the respondent chooses their answer. This provides data which can be easily quantified, but on the other hand does not allow the respondent freedom to deviate from the options presented to them. Because of this, important issues may be missed. On the other hand, structured interviews do allow participants to ask for clarification if there are any questions or items which they do not understand. Due to these characteristics, structured interviews remain a commonly used tool, particularly for market research purposes or opinion polls (Robson, 1993). It was felt, however, that a structured interview would not be appropriate for the present research, as the main aim of the interview study was to elicit new ideas about attitudes towards males and females in sport. It was important that both the respondent and the interviewer could deviate, when appropriate, from the topic being discussed in order to explore alternative lines of thought and for this reason an interview with less structure was required.

Unstructured Interviews

Unstructured interviews are quite different from structured interviews. Whilst both involve 'conversations' in which questions are asked, an unstructured interview does

not have a preset schedule, and the interviewer does not aim to control the type of responses that are given. Instead, the interviewer has an overall research question in mind, and perhaps a few sub-topics with which to 'start the ball rolling'. The respondent is free to say as little or as much as they wish, and the interviewer can follow up different issues with different respondents as they arise. Thus, being able to compare answers from different respondents directly is held to be less important than eliciting rich and meaningful data. These data are usually analysed qualitatively (though content analysis can be used on any interview data to provide quantified results) and together with the flexibility in carrying out the research, means that unstructured interviews are time-consuming (Breakwell, 1995). This type of interview has been used in ethnographic research (e.g. Malinowski, 1989) where the interviewer wishes to 'immerse' themselves in a culture through participating in people's lives, and where formal interviews are not appropriate or would be difficult to initiate (Fontana and Frey, 1994).

Unstructured interviews are relatively time-consuming and this is one reason why this methodology was inappropriate for the present research. As previously mentioned, access to the participants was constrained by the school timetable. It was important, therefore, that the interviews could be carried out fairly quickly but with sufficient leeway for participants to expand on their answers without feeling pressurised. Because unstructured interviews and structured interviews were both considered inappropriate it was decided to employ a semi-structured interview format for Study Four.

Semi-Structured Interviews

Unstructured interviews are time-consuming to carry out and to analyse but they do generate new ideas and explore individuals' social worlds from their own perspectives. Structured interviews, on the other hand, are quick to carry out and provide easily quantifiable data, and do not allow new ideas to be generated. Semi-structured interviews provide a compromise between these two extreme types of interviews, giving a procedure that is quick and straightforward to carry out, as well allowing respondents freedom in their answers.

With semi-structured interviews, the interviewer has a set number of questions, which are asked to all the respondents, but the order in which the questions are asked can vary depending on the way in which the interviewee responds. The interviewer is at liberty to follow up pertinent issues as they arise, and probe respondents to elaborate on their answers. This type of interview allows for the fact that respondents may provide answers to later questions in their responses to other answers (Fielding, 1993). Another advantage of this type of interview is that data collection is more formal and resources of time and money can be more efficiently used than with the open-ended unstructured interview.

Interviews in Attitude and Stereotyping Research

Hayes (1993) suggests that, by analysing interview transcripts and recordings, it is possible to identify key quotes from participants which may reveal their attitudes towards the topic in question. Breakwell (1995) suggests that the research interview is one of the most flexible research tools, and can be used at many stages in the research process. Interviews can be used for identifying topics for further research, for validating other research instruments, for the main data collection, and for confirmation of results from other research tools.

Disadvantages of Interviews

Interviews have been criticised for yielding unreliable results. However, Breakwell (1995) suggests that there is no evidence that interview data are any less reliable or valid than data collected by other methods, and problems with interviews as a tool are common to many other research tools also. In particular, interviewees may still give 'socially desirable' answers. As interviews are usually carried out on a one-to-one basis between the interviewer and the interviewee, this problem may be heightened by the interviewee feeling 'put on the spot'. Reluctance to give truthful and frank answers to sensitive questions may increase if the interviewee feels intimidated by a proximate judge. On the other hand, if an interviewee does not give honest and consistent answers (whether they intend to or not) it is possible in an interview situation to

question any inconsistencies and to clarify any ambiguities that the interviewee may feel regarding the researcher's questions. This is not possible when using a questionnaire methodology. Nevertheless, researcher effects have been well documented, and there are many characteristics of the interviewer and the interview situation which will influence an interviewee's willingness to reveal personal information. For example, interviewees have been found to disclose more information to an interviewer whom they think is more similar to themselves (Breakwell, 1995). Also, Hyman (1954) found interviewer race effects. White respondents gave fewer socially acceptable responses to a white interviewer than black respondents did.

Along with race, consistent effects of sex, age, social class and religion on the interview process (and the data elicited) have all been found (Sudman and Bradburn, 1974). One partial solution to the problem of interviewer effects is to match the interviewer as closely as possible to the characteristics of the research population. There are severe limits to the extent this can be done however, so it is important to take into account the characteristics of the entire interview context, including the interviewer, during the process of analysis (Fielding, 1993).

Interviews can only be an effective research tool if the respondents are able to express themselves verbally. It is particularly problematic to interview children as they may not have the ability to verbalise their thoughts and feelings. It must be noted, though, that other methodologies also rely heavily on communication skills. Questionnaires depend on participants having good literacy skills and being able to understand the questions (Fife-Schaw, 1995).

Interviewer Bias

Concerns have also been raised regarding interviewer bias. Interviews with a lesser degree of structure permit interviewers to change questions, express their own views and challenge respondents to justify their positions. With such freedom there is a danger that an interviewer with a particular point of view may, either intentionally or otherwise (perhaps through non-verbal means), lead a respondent to give answers that are in line with the interviewer's own views. Breakwell (1995) suggests that to

overcome these problems, interviewers should be well-trained to be aware of their own prejudices. If interviewers are conscious of their biases they may be less likely to introduce them during an interview. Again though, this is only a partial solution, and the elimination of bias depends largely on the skill, experience and professionalism of the interviewer (Fielding, 1993). In this respect, interviewing is not so different from any other research tool.

Interviewer bias may also be introduced to the research process during note-taking. Taking notes during an interview is a time-consuming task and there are strong reasons to utilise some form of short cut to reduce the time taken to complete the interview and to maintain the flow of the conversation. However, the interviewer may only write down answers which seem important or relevant at that time, and ignore seemingly insignificant comments which may turn out to be pertinent during analysis. The most common solution to this problem is to use a tape-recorder and then transcribe the interview verbatim. Though the transcription process is particularly time-consuming, the end result is a more accurate and complete record of the interview than notes could provide. A record of the interview also remains for external scrutiny should it be required. A possible problem is that the use of tape-recorders may induce respondent shyness, and restrict the richness of the data collected. Breakwell (1995) however, suggests that recording technology (audio and video) is now so commonplace that there is no reason to suggest that respondents are seriously inhibited.

Indirect Methods of Stereotype and Attitude Assessment

The Link Between Attitudes and Behaviour

Since Katz and Braly's original research in 1933, direct methods of attitude assessment such as the 'Checklist' methodology were widely used for several decades in stereotyping research. Despite possible problems of reactivity there was apparently no difficulty in getting participants to describe the defining traits of racial and ethnic

groups. These methods of assessing attitudes and stereotypes were generally held to be successful.

There was an assumption, though, that there was a direct relationship between holding a stereotype or attitude and prejudicial or discriminatory behaviour. In fact the concept of 'attitude' was developed as a way of predicting future behaviour (Fishbein and Ajzen, 1975). However, this direct link between a person's attitude and their behaviour was called into question by a study by LaPière (1934) who visited over 250 restaurants in the United States accompanied by a Chinese couple and was refused service only once. LaPière then wrote to these same restaurants and asked if they would serve Chinese customers. A large proportion (92%) of the restaurants reported that they would not serve Chinese customers. There are some flaws with LaPière's study. For instance, the couple he travelled with were particularly well-dressed, and this may have impressed the restaurateurs. Also, the people who provided the service during LaPière's visits may not have been the same people who responded to his letter. Nevertheless, this study and many others have highlighted the lack of correlation between what people say and what people do (Wicker, 1969). Thus the attitude-behaviour link became the subject of study and less emphasis was placed on the content of attitudes.

One explanation as to why the link between attitudes and behaviour is so weak is that participants modify their views to give socially desirable responses to attitude measures, whereas their overt behaviour is driven by their 'true' attitude. While reactivity effects are well documented, this explanation gives the impression that everybody lives double lives, saying one thing while doing another. This also calls into question the concept of 'attitude'. After all, if attitudes are so unrelated to behaviour then there would be no reason to assess them. However, there are explanations for the poor correlation between behaviour and attitudes which suggest that 'attitude' can be a useful concept to measure. It would be surprising indeed, if the commonly held assumption that attitudes influence behaviour had no basis in fact whatsoever.

Another explanation for the poor attitude-behaviour correlation is that there may be several competing attitudes which can influence a particular behaviour (Wicker, 1969). For instance, the restaurateurs in LaPière's study may also have held even stronger attitudes against creating a scene. It may be that it was easier for them to ignore one attitude (their prejudice against Chinese people) than it was to create an embarrassing scene.

Fishbein and Ajzen (1975) suggest that the correlation between attitudes and behaviours can be increased by changing the way in which they are measured. One reason for apparent inconsistency is that attitudes are often assessed in general terms, but behaviours are often measured more specifically. Fiske and Taylor (1991) provide the example of an individual agreeing that needy people should be helped through charity (the attitude), but that same individual refusing to donate money to a beggar in the street (the overt behaviour). On the face of it, the behaviour is inconsistent with the attitude. However, agreement with the general statement 'needy people should be helped through charity' does not indicate who should do the helping. An individual may agree with the statement 'The Salvation Army should help needy people through charity' but disagree with the statement 'Individuals should help needy people through charity'.

It is also possible that a specific attitude is measured and then compared to a general behaviour. For instance, asking an individual their feelings towards 'Newcastle United Football Club', and then recording how often they go to all football matches is unlikely to demonstrate a direct link between behaviour and attitude.

Fishbein and Ajzen (1975) also propose that it is important to differentiate between attitudes and intentions. Their theory of reasoned action suggests that a person's overt behaviour can be predicted from their intention to act in a certain way. This intention is only partly a function of their attitudes towards the behaviour in question. Intentions have been found to predict behaviour very well (Fiske and Taylor, 1991) providing that both the intention and the behaviour are measured at the same level of specificity. Also the intention, when measured, must reflect the individual's intention

at the time the behaviour is performed. There are often many intervening factors between the assessment of the intention and the behaviour being performed and this may limit the predictive value of the intention.

While taking these factors into account when trying to predict behaviour has been fairly successful, there are situations where knowing someone's intentions is not as valuable as knowing their expectations (Warshaw and Davis, 1985). Fiske and Taylor (1991) suggest the distinction is important because people can act without having a specific, conscious intention, but with a general expectation. For instance, an individual may expect to drink 10 pints of beer in a week, but not actually have the specific intention to drink 10 pints in a particular week. For the purposes of predicting behaviour the behavioural expectation will be more useful.

Another factor which aids in predicting behaviour from attitudes is the question of whether an individual believes that they are capable of performing the specific behaviour. That is, there is greater consistency between attitudes and behaviour when an individual's self-efficacy is taken into account (Ajzen and Madden, 1986).

Thus the link between attitudes and behaviour is not a direct one; it is not possible to predict an individual's behaviour by knowing their attitudes. However, by measuring attitudes and behaviour at the same level of specificity or generality, by distinguishing between attitudes, intentions and expectations, and by taking into account an individual's beliefs about self-efficacy, predicting behaviour from attitudes can be more accurate.

The problem of social desirability effects and the lack of a consistent attitude-behaviour link led to the development of indirect techniques designed to measure attitudes and stereotypes without the participant being as aware of the subject under investigation. Such techniques are termed 'unobtrusive' or 'non-reactive' methods. In some techniques participants are completely unaware that their attitudes are being measured, while in others they are aware that they are being assessed in some way,

but have no knowledge of the exact nature of the study (Webb, Campbell, Schwartz, Sechrest, and Grove, 1981).

The Lost Letter Technique

In studies utilising Milgram's lost letter technique (Milgram, Mann & Harter, 1965), participants were unaware that they were involved in a study looking at attitudes. The principle behind this method is that it is possible to gauge people's attitudes towards various contentious issues by 'losing' envelopes addressed to organisations involved in that particular issue. It is suggested that members of the public who support a particular cause are more likely to forward the lost letter as they would wish to help that organisation. If they disagree with the cause they would not forward the letter. The 'letters' in stamped envelopes, of course, are placed in public places by the experimenters and the addresses and organisations are fictitious. This technique has been used to try and assess the differences in attitudes across geographical regions by comparing the number of letters returned from different cities.

While the lost letter technique has been hailed as a successful unobtrusive measure of attitudes by some (e.g. Greenwald and Banaji, 1995), this technique has been criticised on the grounds that the return rate for nearly all the studies is around 50% (Dawes and Smith, 1985). If a community is evenly split on an issue then the lost letter technique will appear to be fairly accurate. However, in other circumstances this technique will provide results very different from other measures such as surveys or actual voting percentages. One possible explanation for the return rate always being around 50% is that the technique may tap into people's general politeness or feelings of responsibility rather than their attitude towards specific groups. Furthermore it is assumed that on finding the lost letter individuals will believe that the letter will be of benefit to the organisation. Some people may forward a letter to an organisation that they do not favour in the hope that the letter is a condemnation of the organisation's activities, a withdrawal of subscription, or even a large bill. Rather than trying to second guess the contents, many participants may 'do the right thing' and forward the letter regardless of their own beliefs.

The Wrong Number Technique

A variation on the lost letter technique is the 'wrong number' technique (Gaertner and Bickman, 1972). In this method, the experimenter telephones numbers selected from a 'phone book, and when the call is answered asks if the respondent is "Bob's Garage" (or similar). The 'participant' of course replies that the caller must have the wrong number. The experimenter then explains that they have broken down in their car and wanted to 'phone the garage to arrange collection of the car. However, they have just spent the last of their money in the 'phone, and they ask if the 'participant' would call "Bob's Garage" on their behalf and explain where the car has broken down. The number that the experimenter passes on belongs to a confederate, who records which people do contact the garage and explain the experimenter's predicament. This method has been used to measure racial attitudes (by varying the accent of the experimenter, or by using a stereotypic name) and recording who does make the 'phone call to the garage (e.g. McKenna, 1976).

Other indirect methods can measure attitudes without the participants' direct involvement. For instance, museums have been known to assess the popularity of different exhibits by analysing how often the floor covering in front of the displays needs to be replaced (Webb et al., 1981). The assumption is that the more popular exhibits attract more viewers, and this results in greater wear of carpets. However, this method may give invalid results if an apparently popular exhibit is, say, en route to public lavatories. Another problem is that certain types of people may wear different kinds of footwear which erode the floor covering at different rates. For example, high-heeled shoes are generally more abrasive than flat-soled shoes, and as few men wear high heels, the exhibits appealing to men may be rated as less popular than they really are because men's shoes do less 'damage per visit' than women's shoes.

The advantage of these methods is primarily the unobtrusiveness of the measures. A big disadvantage is the question of validity. It is not clear that these measures actually assess the attitudes they set out to. While all indirect measures suffer from the

problem of validity, face validity may be increased by using techniques where participants know that their attitudes are being assessed but not exactly which ones.

For instance, Asch (1946) presented participants with a list of traits describing a person (intelligent, skilful, industrious, warm, determined, practical, cautious) and in one particular study asked participants to write what they thought about the person. Results showed that participants were able to construct personality sketches of the person, and they generally made inferences about the person beyond the information they had been given. Asch's further work (1952) manipulated the content of the list of traits. Significantly different impressions were elicited by replacing the word 'warm' with 'cold' in the list above. Asch described such traits as being 'central'. This body of research demonstrated that when making impressions of other people, individuals use data (i.e. the list of traits) and theory (i.e. their knowledge of relations between traits). From a methodological perspective, though, this work was important in developing techniques which were higher in validity (participants knew they were making impressions of other people) but did not require participants to know which aspects of the impression were being studied. That is, participants did not know that it was impressions of 'warm' and 'cold' people that Asch was interested in. It was assumed that if participants were unaware of the true nature of the study then reactivity (e.g. social desirability effects) would be eliminated.

A further advantage of Asch's study was that, unlike the wrong number technique or lost letter technique, the material could be presented to participants in a questionnaire format. This makes data collection more straightforward, less time-consuming and less expensive.

The Goldberg Paradigm

As similar technique to that of Asch was employed by Goldberg (1968) in his original work looking at the effect of sex of author on people's evaluations of a piece of work. Just as Asch found radically different impressions by merely changing the words 'warm' and 'cold', Goldberg found that a piece of written work was evaluated in strikingly different ways by changing the name on the front of an article to be

evaluated from 'Joan MacKay' to 'John MacKay' (Goldberg, 1968). The article presented to different participants was identical except for the name of the author. Results showed that females were evaluated less favourably than males, though it should be noted that only female participants were used. It was assumed that the different evaluations were due to gender stereotypes because the articles were identical apart from the sex of the author. While the 'Goldberg paradigm' has been widely used, confounding factors, such as the relative attractiveness of the names used, have led to conflicting and inconsistent results (Swim, Borgida, Maruyama and Myers, 1989; Kasof, 1993).

Whereas the studies based on the Goldberg paradigm are usually questionnaire based, the "Baby X" studies carried out by Seavey, Katz and Zalk (1975) investigated gender stereotyping of infants by asking adults to interact with a 3-month-old baby. In these studies the adults' behaviour was observed and comparisons made across conditions depending on whether the adults had been told the baby was female or male. Again, this manipulation radically changed the way in which the adults reacted to the baby. For example, when the baby was labelled female, participants were more likely to initiate play with a doll. Unlike the restricted choice of responses in questionnaire type studies using a checklist methodology, a study like this allows the participants almost unlimited responses, restricted only by the toys available and time given to the observation.

The four questionnaire studies in the present research were all based on Goldberg's paradigm. As well as the advantages of using a disguised attitude measure, this methodology was chosen for the ease of distribution of materials and for straightforward data collection. While methodologies using 'live' materials, e.g. the Baby X studies, provide interesting results, the amount of time required to set up such studies precluded the use of this kind of methodology in the present research.

Problems with Indirect Methods

Ashmore et al. (1986) highlight two problems with the measurement of stereotypes. One problem is that the different instructions and data analysis across different studies

make comparisons difficult. The second problem is the over-reliance on self-report measures and associated difficulties with reactivity. Ashmore et al. recommend that stereotypes should be measured using a variety of methodologies but that greater attention should be paid to unobtrusive measures.

Dawes and Smith (1980), however, question whether there is a real need for using indirect methods where participants are, to some extent, being misled. Dawes and Smith suggest that these techniques are 'ethically dubious' because participants should be allowed to make an informed decision about whether they want to participate in research. They cannot do this if they are not told the exact nature of the research. Duping or misleading participants should only be considered if it has been established that indirect methods are absolutely necessary. The validity of indirect methods is normally established by comparing their results to results from direct measures. Dawes and Smith suggest that because the results are similar there can be no need to use indirect methods, and it would only be if the results of indirect and direct methods were different that a case for indirect methods could be made. In contrast to this view, Greenwald and Banaji (1995) are quite clear in saying that indirect methods are not just useful in avoiding reactivity from participants, but their use is imperative on methodological and theoretical grounds. Greenwald and Banaji suggest that stereotypes and attitudes largely operate implicitly. That is, most attitudes and stereotypes are not accessible to the introspection required by self-report measures. Amongst the evidence they cite is the fact that it is possible to demonstrate discrimination by people who explicitly disavow prejudice (c.f. LaPière (1934) who found prejudice but not discrimination). Their view is that indirect methods are the only way to measure stereotypes which operate in an implicit fashion.

Summary

Early research investigating stereotyping used overt, direct methods such as the checklist methodology developed by Katz and Braly (1933). Such a direct method suffers from problems with participants reacting to the knowledge that their attitudes and stereotypes are being investigated. Indirect methods of assessing stereotypes

were developed to overcome these problems. In some studies the 'participants' are unaware that their attitudes are being investigated (e.g. Lost Letter technique) while in others they know that they are being asked to make some kind of response (evaluation, judgement, etc.) but they do not know what aspect of their responses is being studied (e.g. Goldberg's paradigm, Baby X studies).

The Present Research

The overall aim of the present research was to investigate gender stereotypes and children's attitudes towards males and females in sport. As previously mentioned the project consisted of five separate but related studies, with the results of each study determining the design and aims of the subsequent study.

Studies One, Two, Three and Five: The Goldberg Paradigm Studies

The aim of Study One was to determine the content of the stereotypes that children hold regarding the sexes in sport. It was hypothesised that stereotypes may have changed since the formative research which was carried out in the 1970s. Because of unique American legislation such as Title IX, it was predicted that stereotypes of males and females in sport in a British population may well be different from those in a North American population, where most research on this topic has been conducted. More specifically, the aim of Study One was to establish which components of traditionally held gender stereotypes currently remained salient components of gender stereotypes in sport for British children.

There are particular problems in assessing the stereotypes and attitudes of children. One of these problems is that children may not have the necessary introspective skills required to answer direct questions regarding their feelings and beliefs. Of course, there are researchers who suggest that even adults do not have the ability to express the stereotypes they hold, because their operation is largely subconscious (Greenwald

and Banaji, 1995). While these claims remain contentious, it is probably the case that children find expressing their stereotypes and attitudes much harder than adults do (Craig, 1992).

The children involved in the present research were aged between 11 and 16 years. Thus, some of them would not have reached the 'formal operational' stage of cognitive development (Piaget, 1929). This stage in cognitive development is characterised by the ability to think abstractly and in general terms. Most children develop this ability around the ages of 12 or 13 years, so there would be a significant section of the sample who would find it difficult to introspect and talk about gender stereotypes in general terms.

To combat these problems, it was decided to utilise an indirect, disguised measure of attitudes based on the 'Goldberg paradigm', in a questionnaire format similar to that used by Wolfson, Ball and James (1985). Each questionnaire was prefaced by a character sketch describing a school pupil, their hobbies and brief biographical details, and then asked the children to describe the character 'based on their first impressions'. The descriptions were elicited by asking the children to endorse a variety of adjectives on five point Likert-type scales ranging from a *very poor description* to a *very good description*. As in Goldberg's original study, the name of the character in the story was varied, so that some children described the character thinking it was a female, while some thought it was a male. This measure of stereotypes is described as 'disguised' in that the children knew that they were being asked about their attitudes and beliefs concerning the character in the story, but because each participant was blind to the other conditions they did not know that the study was specifically concerned with gender stereotypes. It was hoped that socially desirable responses would be avoided.

There are several advantages of this methodology in answering the research question. Firstly, because the children were unaware of the exact nature of the study, socially desirable responses are reduced. It is of course feasible that some children guessed that the study was concerned with gender, so the possibility that the participants gave

socially desirable answers cannot be completely discounted. The second advantage is that by giving the children a specific 'target individual' (i.e. the character in the story) to consider, the task is more 'concrete' and less abstract, and this would hopefully make the task easier to complete. It was hoped that the task involved in completing the questionnaire would be familiar to the children from English classes. Interpretation exercises often ask children to draw conclusions about characters described in short passages of text. It was hoped that children would find it easier to describe a character in a story, than it would be to discuss gender stereotypes in general terms.

Furthermore, the task did not require children to introspect on their own views. They could describe the character without knowing exactly which stereotype was driving that description. This is important in light of Greenwald and Banaji's (1995) assertion that indirect methods are the only way to assess implicitly operating cognitions such as stereotypes.

Another important feature of the questionnaire was the relatively simple response format. The five point Likert scales just asked the children to circle a point on the scale to say whether that adjective was *a very poor description*, *a poor description*, *neither a good nor a poor description*, *a good description*, or *a very good description*. The questionnaire required children to be able to comprehend the story, and to understand the adjectives, but the responses did not rely on the children's own vocabularies or expressive skills. The downside to using these type of scales, though, is that participants are only able to endorse items presented to them.

It must also be noted that this type of response format means that the questionnaire is quick to administer, and can assess the attitudes and stereotypes of a large number of participants in a relatively short time. Keeping the time to complete the task to as short a time as possible also reduces the chances of participants getting bored or losing concentration. Furthermore, by providing quantitative data, the analysis of results is quicker than methods providing data which require to be analysed qualitatively.

A further advantage to the Goldberg paradigm is that it is possible to introduce other variables apart from gender by changing other aspects of the character sketch. For instance, in Study One the character's hobby was also varied. By comparing all the different descriptions it is possible to investigate the interaction effect by manipulating this variable along with manipulating the sex of the story's character.

Study Four: The Semi-structured Interview Study

Studies Two, Three and Five all utilised a similar methodology to that used in Study One. Study Four, however, utilised a different methodology. As will be seen in later chapters, the results of the first three studies, which all used questionnaires based on the Goldberg paradigm, left several questions unanswered. In line with researchers who advocate the use of multiple methodologies in sport psychology research, it was decided then to use a different methodology in an attempt to explicate the results from the first three studies.

Anecdotal evidence (e.g. from school teachers) has suggested that children do hold strong views of males and females in sport, though, as will be seen, the results from the questionnaires did not support this. It was decided to carry out semi-structured interviews, and to perform a qualitative analysis of the data, to test whether some aspect of the questionnaire design had influenced the children's willingness to stereotype the characters they were asked to describe. As mentioned previously, using a different methodology to investigate stereotypes is also in line with the predictions of social judgeability theory (Leyens et al., 1994) and social identity theory (Oakes et al., 1994).

The interview transcripts were analysed qualitatively which involves an "interpretive, naturalistic approach" (Denzin and Lincoln, 1994) to data analysis. This method involves identifying themes and ideas provided by the interviewee, and trying to

answer the research question from the participants' perspectives. Thus, the interviews also allowed the possibility of generating new ideas about children's attitudes towards males and females in sport, whereas, the questionnaires in Studies One, Two and Three only allowed children to agree or disagree with preconstructed views.

Interviews with young children can be particularly problematic (Breakwell, 1995). Some are particularly intimidated when asked questions by an adult and find it difficult to assert themselves and to express their views. Children show an 'acquiescence response bias' i.e. they tend to answer 'yes' to all questions, or agree with all statements regardless of their own views. Another problem is, as Greenwald and Banaji (1995) suggest, that stereotypes operate largely unconsciously so even if a child is keen to talk about their beliefs they may have the problem of not being able to verbalise their attitudes. Interviews, therefore, are only an effective research tool if the interviewee has the skills to express themselves verbally. On the other hand in a semi-structured interview, the children are not constrained in the answers they can give. Interviews provide the opportunity to analyse stereotypes in the children's 'own words'.

Despite the problems mentioned previously, interviews remain a valuable tool for assessing stereotypes and attitudes if steps are taken to reduce children's unease and natural reticence. For instance, questions can be phrased so that a 'yes' or 'no' answer is not required, and interviews can be kept short so that lapses in concentration are minimised. Additionally, social desirability effects can be reduced if the Interviewer does not disclose their own personal views (Breakwell, 1995).

Summary

In recent years there have been more calls for researchers to utilise multiple methodologies to investigate social phenomena. Stereotyping research, in particular, should use different methods to investigate contextual effects on the process of stereotyping. While some researchers suggest that qualitative methods are not compatible with quantitative methods, others suggest that combining methods from

different paradigms is more useful and illuminating than the use of single methods. The present research employed an indirect, questionnaire-based stereotype measure (based on the Goldberg paradigm) which was analysed quantitatively, and a semi-structured interview study which yielded data which were analysed qualitatively.

CHAPTER THREE: THE GOLDBERG PARADIGM (MARK I) : STEVEN AND SUSAN PLAY THE VIOLIN AND GO SWIMMING

Introduction to Study One

The overall aim of the present research was to investigate gender stereotypes in sport. As discussed in Chapter One, the Sports Council (1994b) have identified gender stereotypes as being amongst the barriers preventing females from taking part in sport. However, stereotypes vary over time and between different cultures and it is important to know more about gender stereotypes in sport in the United Kingdom. Knowledge regarding gender stereotypes is limited - particularly knowledge of the stereotypes that children hold. This is despite the fact that it has been established that adults' sport experiences are greatly influenced by their experiences as children. It has also been established that children are particularly influenced by other children - their peer group - in their attitudes towards sport. The present research, therefore, aimed to establish exactly what attitudes, and in particular what stereotypes, children hold of males and females in sport.

The aim of Study One was to explore children's attitudes towards males and females in sport using a disguised attitude measure based on the Goldberg paradigm (Goldberg, 1968). It was hypothesised that a female described as participating in sport would be evaluated in a different way from a female not described as participating in sport, and also in a different way from a male described as taking part in sport. It was predicted that these different evaluations would reflect gender stereotypes of males and females in sport. In particular, it was predicted that these different evaluations of females would reflect negative stereotypes of females participating in sport.

As discussed in Chapter Two, the methodology used for Study One was based on the Goldberg paradigm (Goldberg, 1968), and in particular was similar to the methodology utilised by Wolfson et al. (1985). Goldberg presented participants with a piece of academic writing with the author's name 'incidentally' presented on the front cover, and then asked participants to assess the article for its content. Wolfson et al. (1985), however, presented participants with a short extract about a person, and then asked participants to assess directly the target individual. Similarly, the present study utilised a questionnaire consisting of a short character sketch ('vignette') and the participants were asked to form impressions of that person.

One of the main predictions for the present study was that a female described as taking part in sport would be evaluated differently from a female who is not described as participating in sport. In particular, it was predicted that the sporting female would be evaluated more negatively than the non-sporting female, and that the sporting female would be evaluated more negatively than the sporting male. A problem arises, though, if the sporting female is found to be negatively evaluated in comparison to a non-sporting female. The different evaluation could be attributed to the character being described as taking part in any hobby quite seriously, and not just the fact that they do sport. For this reason, a further level of the 'hobby' variable was introduced. As well as having characters described as playing sport and characters described with no mention of sport, there were also characters which were described as being musicians. In this way, it was possible to compare a female who participates in sport seriously to a female who participates in music seriously. This allows the possibility to test whether any different evaluations of the sporting female were due to sport participation or due to seriously taking part in any hobby.

The participants in Study One were asked to evaluate the characters on a series of trait items. Each item was rated as to whether it was a good description of the

character or not. In this way, participants were able to describe the personality characteristics of the target individual. However, as discussed in Chapter One, the present research has taken a broad definition of stereotypes, in order to include all sorts of beliefs about males and females, and not just the personality traits that typify the sexes. There was a second way in which participants could evaluate the character in Study One, and that was by rating the likelihood of a series of statements explaining why the character in the story had been successful in winning a prize. Studies of attributions for success have shown that females sometimes attribute their successes to unstable factors such as luck (see Chapter One). How females come to adopt such an attributional style is not clear. It may be that significant socialising agents make causal attributions in stereotypic ways and this is then adopted by females in sport. However, it may be the case that observers merely reflect the attributions that individuals themselves make. While it remains problematic to extricate causes and effects within this process, a section on attributions was included to see if the female sports player was, in fact, attributed with different explanations of success than the other characters.

A further aim of Study One was to investigate any differences between older and younger children in the stereotypes they hold. There were three age groups included in this study. The youngest children were those entering secondary education at age 11 years, and the oldest were those who were about to complete compulsory education at age 16 years. A third group of 14-year-olds was also included. This range of ages covers many developmental changes and transitions. As young adolescents approach the end of secondary education they are striving to establish their own identities in terms of sex, race and occupation. Across this time, the influence of the peer group becomes more important and attitudes towards sex-roles may change (Durkin, 1995). It was predicted, therefore, that the older children in this

sample would hold different views regarding the sexes in sport compared to the younger children, reflecting the developmental changes across this age range.

The youngest children selected for this study were in Year 7 (11 and 12 year olds) of secondary school, as it was felt that the large majority of these children would be able to understand the questionnaire and the questions asked of them. In the opinion of the teacher whose class completed the second pilot study, children younger than this may have found some of the words and the rating scale difficult. Another reason for testing only children of this age is that previous research (Bird and Williams, 1980) has found that children younger than 11 years are unable to differentiate between ability attributions and effort attributions.

The oldest children included in this study were 15 and 16-year-olds. This group constitutes the last year of compulsory secondary education. Because schoolchildren over 16 years have chosen to remain in education it is not valid to compare them directly to children who are compelled to stay in education. For this reason, no children aged over 16 years were included in the present research.

As well as investigating the effects of varying the sex and hobby of the character in the story, and the different attitudes across the age groups, the sex of the participant was also an independent variable. Some studies have found that males and females hold similar views of the sexes (e.g. Wolfson et al., 1985; Bird and Williams, 1980), whereas other studies have found that males and females hold different views of the sexes (e.g. Weinberg et al., 1984; Parkhouse and Williams, 1986). It is important to establish whether it is just males who hold negative stereotypes of females or whether these stereotypes are held by everyone. By establishing who holds stereotypes, it may be possible to understand how stereotypes are learned and propagated. Goldberg (1968) concluded from his study that females were prejudiced against females despite

the rise of the Feminist movement. This conclusion would be interesting if Goldberg had used male and female participants in his study. However, Goldberg only used female participants so it is not clear whether males share the prejudiced views that Goldberg found, or even if males held stronger negative views than the female participants.

The aim of Study One was to investigate the attitudes of male and female schoolchildren, from three different age groups, towards male and female characters described as taking part in sport, playing a musical instrument, or not described as doing any sport seriously. It was predicted that the female sport participant would be negatively evaluated or stereotyped in comparison to the other characters.

Method

Design

This study was a $2 \times 3 \times 3 \times 2$ factorial design, varying the sex of the character in the story (Steven Smith or Susan Smith), the hobby of the character (swimming, playing the violin, or no mention of any serious hobby), the age group of the participant (12-year-olds, 14-year-olds, 16-year-olds), and the sex of the participant (male or female). Participants were tested in groups of around 20, with the six different questionnaires distributed randomly within each group. Groups consisted of both males and females, and all participants within a group were from the same school year.

Questionnaire Design

The two main issues concerning the design of the questionnaire were the design of the vignette and the design of the items comprising the dependent measures. Six different

vignettes were constructed by varying the sex of the character (male or female) and the hobby of the character (sport, music or no mention of any serious hobby). An example of a complete questionnaire from Study One can be seen in Appendix A.

Design of Vignette

The overall aim of this study was to test whether a female described as playing a sport was denigrated (i.e. differentially evaluated or stereotyped) in any way compared to a male described in exactly the same way. The aim was also to test whether a female described as playing a sport was denigrated in any way compared to a female described as participating in a non-sport hobby. It was important, therefore, that the sport and hobby chosen were not particularly associated with either sex as this could confound the results. While the perceived appropriateness of various sports for males and females is an interesting avenue of research (see Chapters Four, Five, Six and Seven) the present study aimed to investigate whether a female playing sport *per se* was denigrated in any way.

The philosophy underlying the methodology used in this study is that participants should feel that they have enough information to make an impression of the character in the story. The aim of this type of methodology is to avoid reactions to the knowledge that the study is about gender. It was, therefore, necessary that the vignette did not draw the participants' attention to anything odd or striking. The information that was presented was intended to be straightforward and insignificant so as not to arouse any 'suspicions' regarding the nature of the study. Conversely, the vignette was designed so as not to present participants with such a bland character description that they did not feel able to make an impression of the character.

The aim of the vignette was to generate different evaluations of the character just by varying their sex and by varying the hobby they participated in. The design of the

study required that the vignette sounded 'natural' with both a male character and a female character, and with the different hobbies. It was also important that the vignette did not draw the attention of participants to extraneous features of the character description. Instead, the aim was to give participants enough information to be able to form an impression of the character, but not so much information that the important features (i.e. the character's sex, the hobby played) were made less salient.

In summary, the criteria for designing the vignette were that it should give participants enough information to be able to make an impression, it should not give participants so much information that the factors being studied were overshadowed, it should sound 'natural' in all versions (i.e. with all the different combinations of sex of character and hobbies), and it should not contain any startling or unique information that may confound the other variables or allow participants to guess the nature of the study.

Pilot Study One

The vignette itself described the character's family, the character's favourite school subject (and how they won a prize for this subject), a quote from the character about career ambitions, and the character's hobbies. In order to select appropriate material, a pilot study was carried out with 43 undergraduate students, who were presented with a list of hobbies, school subjects and a variety of quotes concerning future career plans (these lists were constructed by the researcher). The students were asked to indicate whether each item on this list was associated more with males, more with females or was associated equally with males and females. The students were not aware that the purpose of the pilot study was to establish a series of neutral items.

The Vignette

After the first pilot study, the vignette was constructed using the criteria listed above.

The first paragraph of the vignette was as follows:

Steven/Susan Smith is a pupil at a school in Birmingham. He/She has one sister and one brother, who also go to the same school. His/Her father is a fireman with the local fire-brigade and his/her mother is a nurse at a nearby hospital.

The sex of the character in the story was varied by changing the name from 'Steven Smith' to 'Susan Smith' (and all associated pronouns). These names are the same as those used by Wolfson et al. (1985). The character's mother and father were described as having occupations associated with males and females respectively, to make the character appear typical and 'normal'. This meant that any differences in the evaluations of the character could be attributed to the manipulations being employed, and not to other unique features of the vignette. For instance, if the character's mother was described as being an astronaut or a professional golfer, then this may have produced undesired interaction effects on the evaluations of the character

The character was described as having one brother and one sister so that both the male and female characters had one male and one female sibling. The character was described as coming from Birmingham as it is a large city similar to the one in which the present study was carried out, and also because Birmingham encompasses a wide variety of differing socio-economic areas. This deliberately ensured that the social status of the character in the vignette was left ambiguous.

The second paragraph of the vignette was as follows:

At school, Steven/Susan enjoys most subjects, but his/her favourite class is geography. One of Steven/Susan's proudest achievements was receiving

a prize for the best Geography project in his/her school, for which he/she carried out a survey of his/her local area.

Geography was selected from the pilot study as the school subject rated by the students as being the most 'neutral', i.e. not associated more with one sex than the other. The character was described as winning a prize for a project to test whether a female sports player was ascribed with different attributions for success in comparison to a male sports player. As previously mentioned, stereotypes do not just prescribe trait characteristics - they also influence expectations of other people and attributions to explain their behaviour. Females who play sport may be denigrated by having their successes attributed to less 'desirable' causes such as luck. However, to investigate the attributions for success, the domain chosen was a school subject rather than the hobbies. This was because of the problem of establishing equivalent successes at the sport and non-sport hobby. Whereas sport has clear definitions of success and failure, the non-sport hobby (i.e. music) does not.

The third paragraph was as follows:

When asked about what kind of job he/she would like to do when he/she leaves school, Steven/Susan said "I think it is too early to make any decisions about jobs and things - I might go to college or something but I want to keep my options open for now."

This statement was designed to be particularly ambiguous and not to give any indication about the character's likely career. By suggesting 'college or something' it is not established whether the character plans to follow a vocational or academic training course. Again, this statement was rated by the students in the pilot study as

being equally likely to have been uttered by a male or female (i.e. not associated with one sex more than the other).

The final paragraph was the most important as it contained the manipulation of the character's hobby. Research has shown that when participants are presented with a list of new information (e.g. a series of words) the items at the beginning (primacy effect) and end (recency effect) of the list are recalled better than items in the middle of the list (Best, 1995). It was decided to place the information containing the manipulation of the character's hobby last in the vignette to capitalise on any 'recency effect' which may have occurred when the children were reading the vignette.

Outside of school Steven/Susan enjoys going to the cinema, going ten-pin bowling, watching TV, and he/she spends at least two evenings a week reading. [However his/her main hobby is swimming/playing the violin which he/she enjoys so much that he/she trains/practises for an hour and a half every day, and regularly takes part in swimming/music competitions.]

There were three possible variations on this paragraph reflecting the three levels of the manipulation of the character's hobby. All characters were described as going to the cinema, going ten-pin bowling, watching TV and reading. All these activities were rated by the pilot study participants as not associated with one sex more than the other. After this sentence the character was either described as a sportsperson (a swimmer), as a musician (a violinist), or no mention was made of any hobby to which the character was especially committed (and the sentence in parentheses was omitted completely). Swimming and playing the violin were again selected from the pilot study as being 'neutral' activities. The character was described as spending two evenings a week reading so that participants in all conditions could answer the 'manipulation check' question (Question 3, see below) which asked 'What do you

think Steven/Susan's main hobby is?' and 'How much time does he/she spend doing it?'. In this way, all of the items in all conditions were identical except for the factors being manipulated. Maintaining equivalence across the conditions avoided the possibility that the results could be influenced by small differences which were not related to the factors being studied.

An example of the complete vignette is as follows:

Steven Smith is a pupil at a school in Birmingham. He has one sister and one brother, who also go to the same school. His father is a fireman with the local fire-brigade and his mother is a nurse at a nearby hospital.

At school, Steven enjoys most subjects, but his favourite class is geography. One of Steven's proudest achievements was receiving a prize for the best geography project in his school, for which he carried out a survey of his local area.

When asked about what kind of job he would like to do when he leaves school, Steven said "I think it is too early to make any decisions about jobs and things - I might go to college or something but I want to keep my options open for now."

Outside of school Steven enjoys going to the cinema, going ten-pin bowling, watching TV, and he spends at least two evenings a week reading. However his main hobby is swimming which he enjoys so much that he trains for an hour and a half every day, and regularly takes part in swimming competitions.

Item Design

The items on this questionnaire were in two main sections. One asked each participant to rate the character from the story on a series of traits (Trait items). The other section consisted of a variety of possible explanations of why the character has been successful in winning the prize for the geography project (Attribution items).

Trait Items

The main criterion for the list of trait items was that there should be a wide variety of traits which would be likely to elicit gender stereotypes. It was also important that the youngest children in the sample (11 year olds) should understand all the traits as well as the vignette. The trait items in the present study were derived from the Bem Sex Role Inventory (Bem, 1974). While this inventory has been criticised in the way it has been employed in research (see Chapter One) it was felt that this inventory would provide a valid starting-point for adjectives that would elicit different evaluations of the male and female character.

Pilot Study Two

A second pilot study was carried out with 21 eleven-year-old children to test whether they understood the task required of them and that they understood the vignette and the items. Table 3.1 shows the items that were selected for the present study and the original Bem items (where appropriate) from which they were derived, as well as the original Bem classification of whether the item contributes to the masculine or feminine scales. From the pilot to the main study, only one adjective - 'ambitious' - was eliminated. This was due to a high proportion of pilot study participants not understanding the word or not answering that particular item. Items marked with an asterisk in Table 3.1 were not included in the pilot study but were added later to give a total of 23 items, consisting of 8 masculine, 8 feminine, and 7 neutral traits, according to Bem's classification. Some items (e.g. 'loves children') were amended

from those in the BSRI ('love children') so that the new items made sense when used to describe a third person rather than the first person as required by the BSRI.

Present Study	Original Item in BSRI	BSRI classification
Caring	Compassionate	feminine
Works hard	Conscientious	neutral
Competitive	Competitive	masculine
Loves children	Love children	feminine
Confident	Assertive	masculine
Makes decisions easily	Make decisions easily	masculine
Aggressive*	Aggressive	masculine
Popular*	Likeable	neutral
Easily fooled*	Gullible	feminine
Shy	Shy	feminine
Reliable*	Reliable	neutral
Athletic	Athletic	masculine
Would be a good leader	Have leadership abilities	masculine
Sensitive to the needs of others*	Sensitive to needs of others	feminine
Independent*	Independent	masculine
Gentle	Gentle	feminine
Lazy	Inefficient	neutral
Copes well on own	Self-sufficient	masculine
Loyal	Loyal	feminine
Willing to take risks*	Willing to take risks	masculine
Cheerful	Cheerful	feminine
Honest	Truthful/Sincere	neutral
Jealous of other people	Jealous	neutral

Table 3.1 Traits Items, Original BSRI items and BSRI Classification (*items added after pilot study)

Other items were changed in consultation with the teacher of the class involved in the pilot study. These changes were intended to make the words more comprehensible to as many of the children as possible. For instance, the class teacher felt that

'conscientious' would not be understood by most of the children in her class, so this item was changed to 'works hard'. By changing items in this way, there may have been a change in the exact meaning of the items. However, the intention was to produce a series of items which would be likely to elicit different evaluations of males and females. It was not important to maintain the substance of the BSRI. It was important, however, to maintain roughly similar numbers of 'masculine' and 'feminine' items and the original BSRI classification was utilised to this end. The inclusion of 'neutral' items was important to provide a wide variety of traits and also to mask the true nature of the study. Had only 'gender salient' traits been used, the participants may have guessed the aims of the research. The items are presented in the table in the same order as they were presented in the questionnaire.

Attribution Items

As with the trait items, the section of attribution items was intended to cover a variety of possible explanations of why the character had won a prize in geography. A series of eleven possible explanations was constructed using the dimensions identified by Weiner (1985) of 'stability' and 'locus of control'. Stability refers to whether the cause is believed to be stable across time or unstable. Locus of control refers to whether the cause is perceived to lie within a person (i.e. an internal cause) or is external. For example, the explanation 'Steven is good at geography' suggests that the reason Steven won the geography prize can be attributed to an internal and stable cause, i.e. his ability. Table 3.2 shows the explanations which were presented to participants as well as the classification on the two causal dimensions.

The causal classifications presented in Table 3.2 are not definitive. Different people may classify the same explanation in different ways, even on the same dimension. For example, for some people 'luck' may lie internally, whereas for others 'luck' lies externally. Also, it is possible to classify any given explanation on many different

dimensions and not just on locus of control or stability. For the purposes of this study, the intention was only to present participants with a variety of different possible explanations to allow for different evaluations of the characters. For each of these attribution items participants were asked to rate how good the explanation was on a five point Likert-type scale, ranging from 1 (*very poor explanation*) to 5 (*very good explanation*).

	Locus of control (internal or external)	Stability (stable or unstable)
Steven tried very hard	internal	unstable
Steven was lucky	external	unstable
Geography is not a difficult subject	external	stable
Steven is good at geography	internal	stable
Steven is a hard worker	internal	stable
Steven was very fortunate on this project	external	unstable
Steven's parents encouraged him to work hard on this project	external	unstable
Steven finds geography an easy subject	internal	stable
Steven has been brought up to put maximum effort into his schoolwork	external	stable
Steven is just a lucky kind of person	internal	stable
Steven has always had a natural talent for geography	internal	stable

Note : For conditions with female character, 'Steven' was replaced with 'Susan'. All appropriate pronouns were also changed.

Table 3.2 Attribution Items and Classification on Stability and Locus of Control Dimensions

In the second pilot study, a section was included which asked the participants to generate their own explanations. The children generally found this section difficult to complete and did not generate any new causes that were not already included in the section of explanations listed above. This section also considerably added to the time taken to complete the questionnaire so this open-ended section was eliminated from the main study.

A further problem which arose from the second pilot study was that some children may have been completing the questionnaire by circling various numbers without actually understanding the task. In the opinion of the class teacher, a few children were such poor readers that it is unlikely that they understood the vignette and so they had probably just randomly circled the numbers next to the items. To check that each participant did understand the vignette an additional set of questions was included in the questionnaire after the vignette. This asked factual questions about the character and their hobbies. As the participants were able to refer back to the story at any time when filling in the questionnaire, it was possible to eliminate from the sample used for analysis, any participants who were unable to answer these questions fairly accurately.

The questions asked were as follows:

1. How many people are there in Steven's family, including Steven?
2. What does Steven's father do?
3. a) What do you think Steven's main hobby is?
 b) How much time does he spend doing it?
4. What subject does Steven like best at school?

The main criterion for a participant to be included in the analysis was that they answered both parts of Question 3 correctly. As this question referred to the factors being manipulated it was decided that if the participant was unable to answer it

correctly, then the rationale behind the questionnaire would be undermined. The only other criterion was that participants answer at least one other question correctly. Out of 301 completed questionnaires 16 were eliminated from analysis on the basis that the children did not understand the vignette well enough.

Instructions to Participants

The written instructions to participants prefixing the trait items were as follows:

Now, although you only know a little about Steven, I want you to imagine what kind of person you think he is. Read each of the words and phrases in this list and then indicate how well you think that word or phrase describes Steven.

Circle the number 1 if you think the word is a very poor description of Steven

Circle the number 2 if you think the word is a fairly poor description of Steven

Circle the number 3 if you think the word is neither a good nor poor description of Steven

Circle the number 4 if you think the word is a fairly good description of Steven

Circle the number 5 if you think the word is a very good description of Steven

If there is a word or phrase which you do not understand write a cross next to it, and move onto the next one.

The list of trait items then followed. Next to each item were the numbers 1 to 5 across the page.

For the attribution items the written instructions were as follows:

In the story Steven won a prize for his geography project. Can you think of some reasons why he won the prize?

Below are some possible explanations as to why Steven won the prize. You may think that some are more likely explanations than others. Try and imagine what kind of person Steven is, and then indicate how likely or unlikely you think that explanation is to be true. Circle only one number for each explanation.

Circle the number 1 if you think it is a very unlikely explanation

Circle the number 2 if you think it is an unlikely explanation

Circle the number 3 if you think it is neither a likely nor unlikely explanation

Circle the number 4 if you think it is a likely explanation

Circle the number 5 if you think it is a very likely explanation

The list of attribution items then followed down the page. Next to each item were the numbers 1 to 5 across the page.

After the section of attribution items, participants were asked to write their age and sex. They were also asked to write down the hobbies that they spent most time on,

and to indicate how much time they spent on each. It was hoped that the questionnaires of sport participants could be compared to those of non-sports participants. However, due to large differences in the ways in which the children interpreted this section (confounded by a lack of time to explain to them what was required) and subsequent difficulties in categorising the children, it was not possible to carry out such an analysis.

Participants

The participants were 301 schoolchildren in years 7, 9 and 11 in a comprehensive school in the North East of England. The mean ages for each year group were 11.9 years ($SD=0.40$), 13.7 years ($SD=0.45$), and 15.8 years ($SD=0.45$) respectively.

Procedure

The questionnaires were distributed by the researcher to groups of approximately 20 schoolchildren. The class teacher was present throughout each session. The teacher was asked to seat the children as if they were under 'exam conditions' so that the children would not be able to read each other's questionnaires and so compromise the principles behind the methodology i.e. that participants must be blind to each of the other conditions. When introducing the researcher, the teacher was asked to make no reference to 'sport', 'gender', 'sex', 'stereotypes' or 'psychology'. This was again to ensure that the children did not guess the nature of the study.

Participants were told that the researcher was from the University and was carrying out a project looking at how people make first impressions of other people and that

they were being asked to help in this project by completing a questionnaire. The questionnaires were distributed face down while the researcher explained the procedure. It was emphasised that there were no right or wrong answers and that, despite the seating arrangements, it was not a test or exam. Nevertheless, it was vital that the participants concentrate on their own questionnaire, as it was their individual opinions and views which were important. The children were told that if they did not understand anything they were simply to place a cross next to the word or item and carry on. The researcher explained to the children how to complete the five-point scales and then asked the children to work at their own pace through the questionnaire.

The teacher was asked not to help the children by explaining the meanings of words. The participants were asked to raise their hands if they had any problems. If their query was concerning the content of the vignette or the meaning of any words they were told just not to worry about it and to carry on. It was important that the researcher or teacher did not induce the children to interpret the questionnaire in a particular way.

Completing the questionnaire took between 15 and 25 minutes depending on the reading ability of each participant. With the approval of the teacher in charge, the participants were not debriefed after the questionnaires had been completed unless they specifically asked. This was because other groups in the same school were still to be tested at later times and it was important that these other groups did not have prior knowledge of the study. It was not possible to guarantee that the children did not deduce the nature of the study and then talk to other children about it. However, by not debriefing participants it was hoped that participants in later groups would not be forewarned. Any children who did query the nature of the study were informed that the researcher was interested in the sports that boys and girls do, and what they

think about them. After all the data collection had been completed the school was provided with feedback on the study.

Results

Individual items were analysed using a 4-way factorial anova. Contrary to expectations, across all the dependent measures, no consistent main effects of sex of character and no significant interactions between the variables of sex of character and hobby of character were found. The results demonstrated quite clearly that the male and female characters were described in very similar ways, and that varying the hobby did not create any differences between the descriptions of the male and female characters. Also, there were no consistent significant main effects or interactions of sex of participant or of age of participant.

To illustrate the general pattern of results, Figure 3.1 shows the mean ratings for 'aggressive' for the male and female characters across the different hobbies. 'Aggressive' is a trait which might be expected to elicit different evaluations for the male and female characters. The Sports Council (1994b) suggests that females are prevented from taking part in sport through negative stereotyping. A possible mechanism by which negative stereotypes operate, it could be argued, is that females do not take part in sport because they are not aggressive enough - males are aggressive and this is why they take part in sport. For the item 'aggressive', the male swimmer had a mean rating of 1.86 (SD=1.05) on the five point scale with 5 being a *very good description* and 1 being a *very poor description*. The female swimmer had a mean rating of 1.92 (SD=1.20), the male violinist had a mean rating of 1.83 (SD=0.97) and the female violinist had a mean rating of 1.67 (SD=1.02). For the condition where no serious hobby was mentioned the male character had a mean rating of 1.96 (SD=1.02) and the female character had a mean rating of 1.56 (SD=0.97). Two-way factorial analysis of variance showed no significant interaction between the variables sex of character and hobby of character, $F(2,242) = 0.807, p > 0.05$.

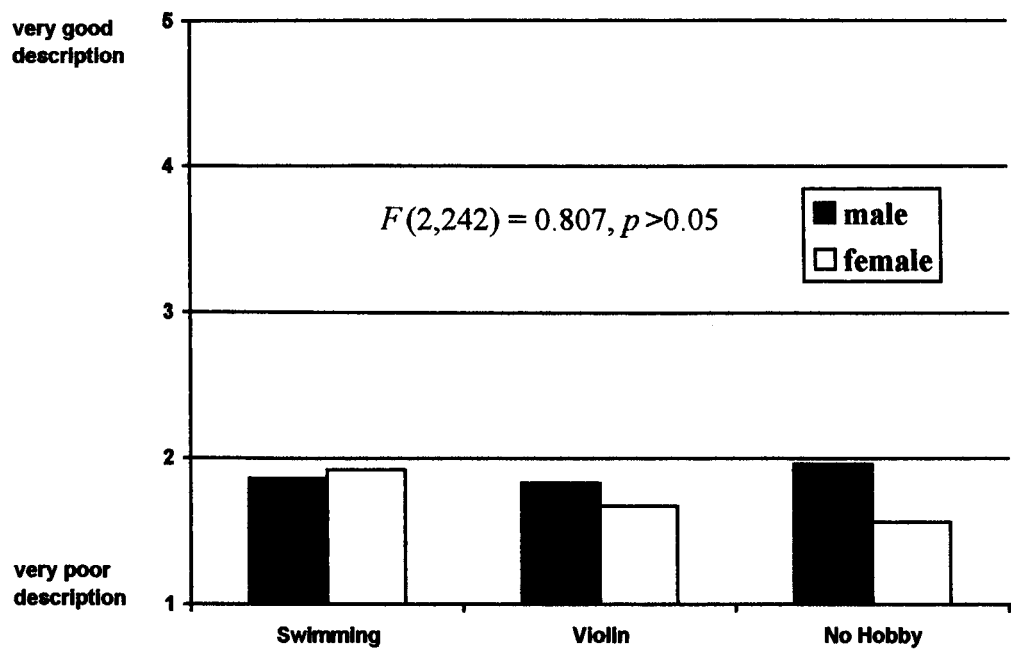


Figure 3.1 Mean ratings for 'Aggressive' for Sex of character by Hobby of character

The mean ratings for the rest of the trait items echoed the results for 'aggressive'. Table 3.3 shows the mean ratings (and standard deviation and number of participants in each condition) for all the trait items for the male and female characters across the different hobbies. Shaded boxes in Table 3.3 indicate results discussed in the text.

Similar results were found for the attribution items. The male and female characters were not attributed with different explanations for their success regardless of whether they were described as doing a sport or not, and again there were no main effects or interactions for sex of participant or age of participant. Table 3.4 shows the mean ratings of each explanation (and standard deviation and number of participants) for the male and female characters across the different hobbies. The shaded boxes in Table 3.4 indicate results that are discussed in the text.

	Male character			Female character		
	Swimming	Violin	No hobby	Swimming	Violin	No hobby
Caring	3.30 (0.81) n=46	3.30 (0.91) n=50	3.17 (0.71) n=46	3.24 (1.09) n=49	3.17 (1.14) n=48	3.22 (0.97) n=45
Works hard	4.70 (0.59) n=46	4.58 (0.73) n=50	4.71 (0.55) n=45	4.76 (0.66) n=50	4.63 (0.61) n=48	4.67 (0.52) n=45
Competitive	4.33 (0.73) n=46	3.88 (0.91) n=48	3.91 (0.75) n=43	4.15 (0.91) n=47	3.87 (1.13) n=47	3.82 (1.05) n=45
Loves children	2.63 (0.90) n=46	2.44 (1.07) n=50	2.72 (0.91) n=46	2.59 (1.10) n=49	2.49 (0.91) n=47	2.73 (1.01) n=45
Confident	3.96 (0.93) n=45	4.14 (0.73) n=50	4.02 (0.71) n=46	3.90 (0.92) n=49	4.10 (0.83) n=48	3.93 (0.84) n=45
Makes decisions easily	2.78 (1.21) n=46	2.78 (1.23) n=50	2.85 (1.28) n=46	2.80 (1.19) n=49	2.93 (1.22) n=46	2.71 (1.08) n=45
Aggressive	1.86 (1.05) n=44	1.83 (0.97) n=48	1.96 (1.02) n=45	1.92 (1.20) n=48	1.67 (1.02) n=48	1.56 (0.97) n=45
Popular	3.22 (0.81) n=46	2.86 (0.99) n=50	3.04 (1.03) n=46	3.04 (0.97) n=46	3.19 (0.98) n=48	3.09 (1.02) n=45
Easily fooled	1.76 (0.90) n=46	2.02 (1.03) n=49	2.02 (1.06) n=45	1.78 (1.06) n=45	1.90 (1.06) n=48	1.73 (0.94) n=45
Shy	2.38 (0.91) n=45	2.55 (1.14) n=47	2.50 (1.21) n=46	2.41 (1.18) n=46	2.65 (1.00) n=48	2.60 (1.18) n=45
Reliable	4.00 (0.88) n=45	4.04 (0.82) n=48	3.98 (0.77) n=46	3.98 (1.04) n=46	4.00 (1.05) n=48	3.93 (1.10) n=45
Athletic	4.87 (0.40) n=45	2.59 (1.17) n=49	2.80 (1.17) n=46	4.68 (0.84) n=50	2.68 (1.14) n=47	3.07 (1.23) n=45

Cell contents : mean , standard deviation, number of participants

(rating scale: 1= very poor description, 5 = very good description)

Table 3.3 Mean ratings of trait items for Sex of character by Hobby of character
(continued over)

	Male character			Female character		
	Swimming	Violin	No hobby	Swimming	Violin	No hobby
Would be a good leader	4.02 (1.13) n=46	3.56 (1.21) n=50	3.83 (1.00) n=46	3.68 (1.06) n=50	3.63 (1.06) n=48	3.71 (1.10) n=45
Sensitive to the needs of others	3.37 (0.88) n=46	3.06 (0.90) n=49	3.20 (0.89) n=45	3.16 (0.96) n=50	3.27 (1.03) n=48	3.27 (0.99) n=45
Independent	3.91 (0.76) n=46	3.71 (1.04) n=49	3.80 (0.94) n=45	3.70 (1.14) n=47	3.96 (0.98) n=47	3.66 (0.89) n=44
Gentle	3.36 (1.01) n=44	3.44 (0.95) n=50	3.36 (0.88) n=45	3.52 (1.16) n=50	3.65 (1.14) n=48	3.41 (1.02) n=44
Lazy	1.18 (0.44) n=45	1.84 (1.13) n=50	1.84 (1.24) n=45	1.47 (1.12) n=49	1.56 (1.03) n=48	1.67 (1.00) n=45
Copes well on own	3.98 (0.91) n=46	4.34 (0.89) n=50	4.15 (0.92) n=46	3.96 (1.11) n=50	4.13 (0.84) n=48	4.33 (0.90) n=45
Loyal	3.58 (0.92) n=45	3.62 (0.88) n=50	3.70 (0.89) n=46	3.57 (0.84) n=49	3.69 (1.01) n=48	3.47 (0.81) n=45
Willing to take risks	3.07 (1.10) n=46	2.62 (1.03) n=50	3.02 (0.99) n=45	3.04 (1.25) n=48	2.92 (1.20) n=48	2.71 (1.04) n=45
Cheerful	4.09 (0.94) n=46	3.66 (0.89) n=50	3.80 (0.81) n=46	3.88 (0.90) n=49	3.77 (1.15) n=48	3.80 (0.87) n=45
Honest	4.26 (0.83) n=46	3.98 (1.02) n=50	3.96 (0.94) n=46	3.98 (1.15) n=49	4.13 (1.00) n=48	3.96 (1.07) n=45
Jealous of other people	1.50 (0.89) n=46	1.76 (1.08) n=50	1.48 (0.84) n=46	1.76 (1.20) n=49	1.54 (0.90) n=48	1.55 (0.66) n=44

Cell contents : mean , standard deviation, number of participants

(rating scale: 1= very poor description, 5 = very good description)

Table 3.3 (contd.) Mean ratings of trait items for Sex of character by Hobby of character

	Male character			Female character		
	swimming	violin	none	swimming	violin	none
S. tried very hard	4.78 (0.55) n=46	4.90 (0.36) n=50	4.74 (0.53) n=46	4.76 (0.66) n=50	4.81 (0.53) n=48	4.91 (0.29) n=45
S. was lucky	1.78 (1.15) n=46	1.94 (1.11) n=50	1.93 (1.02) n=46	1.98 (1.06) n=50	2.02 (1.21) n=48	2.04 (1.00) n=45
Geography is not a difficult subject	2.46 (1.11) n=46	2.54 (1.36) n=50	2.48 (1.13) n=46	2.82 (1.24) n=50	2.53 (1.18) n=47	2.47 (1.14) n=45
S. is good at geography	4.70 (0.55) n=46	4.74 (0.44) n=50	4.70 (0.59) n=46	4.73 (0.49) n=49	4.71 (0.54) n=48	4.58 (0.62) n=45
S. is a hard worker	4.71 (0.46) n=45	4.70 (0.61) n=50	4.70 (0.70) n=46	4.76 (0.48) n=49	4.81 (0.39) n=48	4.62 (0.61) n=45
S. was very fortunate on this project	2.13 (0.92) n=45	2.50 (1.16) n=50	2.82 (0.94) n=45	2.86 (1.26) n=50	2.31 (1.03) n=48	2.36 (1.05) n=45
S.'s parents encouraged him/her to work hard on this project	3.22 (1.11) n=46	3.06 (1.02) n=50	3.39 (1.14) n=46	3.20 (1.18) n=50	3.21 (1.17) n=48	3.33 (1.04) n=45
S. finds geography an easy subject	3.78 (0.87) n=46	3.71 (0.96) n=49	3.93 (0.88) n=46	3.80 (1.18) n=50	3.73 (1.09) n=48	3.76 (0.86) n=45
S. has been brought up to put maximum effort into his/her schoolwork	3.46 (1.09) n=46	3.42 (0.86) n=50	3.87 (0.93) n=46	3.68 (1.06) n=50	3.65 (1.02) n=48	3.71 (0.99) n=45
S. is just a lucky kind of person	1.70 (0.89) n=46	1.84 (0.83) n=49	1.85 (0.97) n=46	1.88 (1.13) n=49	1.71 (0.82) n=48	1.87 (1.01) n=45
S. has always had a natural talent for geography	3.48 (1.05) n=46	3.64 (1.12) n=50	3.76 (0.95) n=46	3.76 (1.18) n=49	3.60 (1.07) n=48	3.60 (1.03) n=45

Cell contents : mean , standard deviation, number of participants
(rating scale: 1= very poor description, 5 = very good description)

Table 3.4 Mean Ratings for Attribution Items for Sex of Character by Hobby of Character.

These results provided no evidence that the female character who was described as being a serious sports participant was denigrated, stereotyped or negatively evaluated in comparison to the male characters or other female characters. However, while no gender stereotyping was apparent, there is evidence to suggest that the participants evaluated the characters in line with stereotypes of sport. For instance, the character who was described as a swimmer (regardless of sex), was seen as more athletic (see Figure 3.2) and more competitive (see Figure 3.3) than the violinist and the character who was not described as having a serious hobby. The swimmers had a mean rating for 'athletic' of 4.77 (SD=0.86, n=95), the violinists had a mean rating of 2.64 (SD=1.15, n=96), and the no-hobby characters had a mean rating of 2.93 (SD=1.20, n=91). An analysis of variance showed a significant main effect for the hobby of the character, $F(2,246) = 124.99, p < 0.001$. A Scheffé post-hoc analysis revealed that the swimmers were rated as significantly more 'athletic' than the other characters.

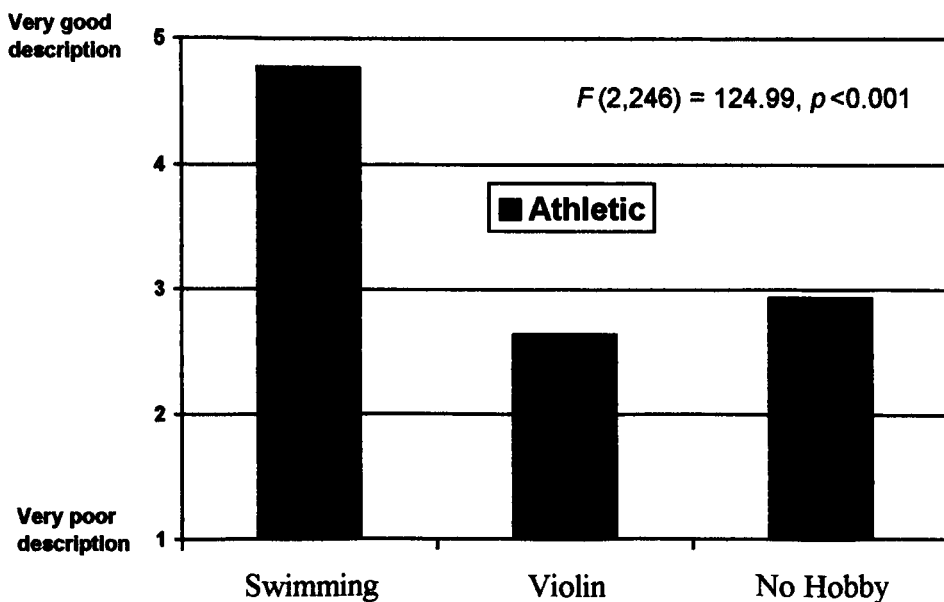


Figure 3.2 Mean ratings for *Athletic* by Character's Hobby

Similarly, for 'competitive' (see Figure 3.3), the swimmers had a mean rating of 4.24 (SD=0.83, n=93), the violinists had a mean rating of 3.87 (SD=1.02, n=95), and the no-hobby characters had a mean rating of 3.86 (SD=0.91,n=88). An analysis of variance showed a significant main effect for the hobby of the character, $F(2,240) = 4.78, p<0.01$. Again, a Scheffé post-hoc analysis showed that this main effect was due to the swimmers being rated significantly more 'competitive' than the other characters.

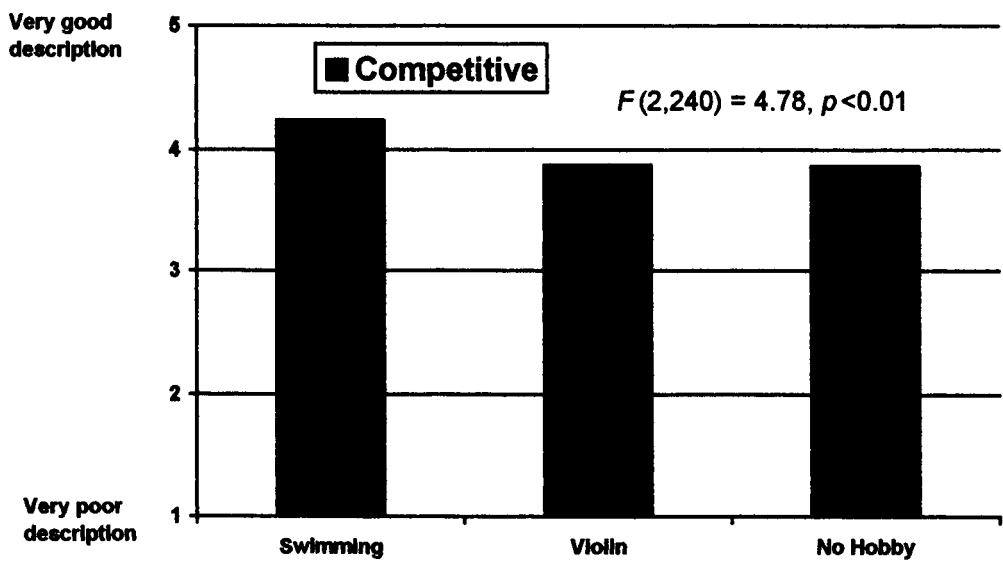


Figure 3.3 Mean ratings for *Competitive* by Character's Hobby

Discussion

In summary, the results provide no evidence of gender stereotyping or that the female sport-playing character was evaluated in a different or negative way in comparison to the male sport-playing character. Nor is there any evidence to suggest that the female

swimmer was evaluated in a different or negative way in comparison to the female violinist or the female character with no-hobby. There is, however, evidence that the participants evaluated the characters in line with stereotypic views of sports and music, in that the swimmer was rated as being more athletic and competitive than the violinist and the no-hobby character. There were no consistent significant effects of the sex of participant or of the age of participant.

Contrary to expectations, the female sports player was not evaluated in a negative way compared to the other characters. One possible explanation for these results is that the children did not understand the questionnaire or that they did not follow the principles behind the five-point Likert-type scales. However, the fact that the children described the sports player in line with stereotypic views suggests that they were able to picture the character, determine what they might be like, and then describe the character using the scales. While finding that the swimmer is perceived as being more competitive and more athletic than a violinist is not entirely surprising, it does confirm that the children understood the questionnaire. It is interesting to note that in the vignette the violinist was described as being 'competitive'. The last line of the vignettes in the music conditions read as follows:

However his/her main hobby is playing the violin which he/she enjoys so much that he/she practises for an hour and a half every day, and regularly takes part in music competitions.

The children must, therefore, have been using stereotypes of swimmers in order to rate them as being more competitive. In other words, the children did not just rely on the information given to them in the story. Again this suggests that while the methodology was not successful in eliciting *gender* stereotypes it was successful in eliciting *other* stereotypes.

The fact that there were no significant effects for sex of participant may not be surprising. Previous research, using similar types of methodology, has produced conflicting results with some studies finding that males and females perceived the sexes differently, while others have found similar views from males and females. For example, Wolfson et al. (1985) found that male and female participants had similar attitudes towards the vignette characters. However, Weinberg et al. (1984) found that male and female participants differed in their attitudes towards a hypothetical female character. In the present study, although the male and female participants quite clearly described the characters in similar ways, this might not have been the case if the questionnaire had been successful in eliciting gender stereotypes.

Contrary to expectations, no consistent age differences were found. It was considered that children just starting secondary education might well have different attitudes towards the sexes than those just finishing compulsory secondary education, especially as during puberty the differences between the sexes become much more salient. Ruble and Ruble (1982) suggest that "...the nature of sex-typing changes during adolescence, in accordance with a newly emerging identity as a sexual being and strong peer pressures". However, as with the sex of participant, just because no age differences were found does not mean that they would not exist had the questionnaire elicited gender stereotypes.

It is surprising that there were no consistent significant effects of sex of character, and no significant interactions between sex of character and the hobby of the character. Previous research using the vignette method (Wolfson et al., 1985; Weinberg et al., 1984; Parkhouse and Williams, 1986) has found differences in the attitudes towards males and females in sport. Wolfson et al. (1985), in a study using undergraduate students as participants, manipulated the sex of the character and their motivation for taking part in sport. Strong interactions were found between these two factors. In

particular, the male was found to be denigrated if he was described as taking part in sport to improve his appearance. The female character tended to be evaluated more positively if she was described as taking part in sport to improve her appearance. However, in common with the present study, it was found that both male and female participants held similar views.

Weinberg et al. (1984) presented male and female athletes with descriptions of a new coach whose sex was varied across conditions. Male and female athletes were found to hold similar attitudes towards the hypothetical male coach, but males displayed more negative attitudes towards the female coach than the females athletes did. In another study looking at attitudes towards male and female coaches, Parkhouse and Williams (1986) found that male and female athletes rated male coaches more positively than female coaches, and male athletes rated the female coach more negatively than the female participants.

The results of the present study indicate that the children did understand the questionnaire and that the manipulation was at least partially successful, as there were significant and expected effects of varying the hobby of the character in the vignette. However, because no gender stereotypes were elicited, it is not possible to conclude that the children in this sample do not hold stereotypic views of the sexes in sport. While this may be a possibility, such a conclusion would contradict the findings of many other researchers and the conclusions drawn by the Sports Council regarding female participation in sport. On reflection, there are certain features of the questionnaire used in the present study which could possibly explain the results without concluding that children hold non-stereotypic views of the sexes.

The reason that background information was included in the vignette was to give readers the impression that they knew a lot about the character so that they could

make an impression of that character. In fact, the aim was to provide such 'typical' information (i.e. information that could be applied to almost anybody) that the children could only resort to stereotypes to make impressions of the characters. It was also important that items in the background information were not associated with one sex more than the other. The items used in the vignette (the pastimes, quote and geography) were those items judged in a pilot study by undergraduates as not being associated with one sex more than the other. However, it may be the case that what was considered typical by the researcher was not considered typical by the children in the sample. The fact that the character was described as winning a geography prize (in order to look at attributions for success) may in fact have made the character appear hard-working and studious. The character was also described as saying that they would consider going on to college after finishing school which again may have indicated a hard working image. Furthermore, 'spending at least two evenings a week reading' may appear innocuous on its own, but in conjunction with the other items it may have reinforced a studious image, which then overshadowed any effect of varying the character's sex. Evidence to support this possibility is the fact that, for both the male and female characters, the attributions which were based on ability and effort were rated as being much more likely than those based on luck. Similarly, the trait item 'works hard' was rated as a good descriptor of all of the characters (means ranged from 4.58 to 4.76), and the item 'lazy' was rated as a poor descriptor of all of the characters (means ranged from 1.18 to 1.84). Thus, the manipulation of varying the character's hobby may have been overshadowed by the portrayal of all the characters as being hard working and studious.

In a similar vein, the amount of background information presented in the vignette may also have influenced the results. While this information was identical across all the conditions, it may be that there was enough information for the participants to

describe the characters without resorting to using gender stereotypes. Locksley, Borgida, Brekke and Hepburn (1980:p830) point out that

Social stereotypes may affect judgements of individuals about whom little else is known besides their social category. But as soon as individuating, subjectively diagnostic characteristics of a person are known, stereotypes may have a minimal, if any, impact on judgements of that person.

It may be that participants did not categorise the characters into 'male sportsperson' or 'female sportsperson' because there was much more information available to them than the characters' social categories. On the other hand, the information in the vignette is unlikely to have been 'diagnostic' in that most of the traits items were not referred to in the vignette. That is, there was nothing in the story to indicate whether the character was 'loyal' or 'cheerful'. To be able to rate the item, the participants had to infer that information from the impressions that they had formed of the character. While the vignette may certainly have provided the 'answers' to some of the traits and attribution items, the information provided to the participants was by no means diagnostic for all of them.

The choice of sport (i.e. swimming) may also have affected the subsequent ratings. While sport in general is seen as a masculine domain, swimming is seen as appropriate for both sexes (Cszima, Wittig and Schurr, 1986). Swimming was chosen after the pilot study as the sport to be used because the aim of the present study was to elicit evaluations of a female playing sport *per se*. This means that this questionnaire was not necessarily testing the attitudes towards females in a 'masculine' domain. Perhaps, therefore, it should not be surprising that males and females are viewed in a similar

manner when they are taking part in a neutral sport. It is almost certain that attitudes towards females playing an 'appropriate' sport are going to be different from those towards a female playing an 'inappropriate' sport.

There was also a problem in using the BSRI as the source of traits on which the characters were rated. Bem's inventory only includes positive traits which were associated with one sex more than the other and no negative traits (Although 'gullible' is not necessarily a positive trait it was rated as being more desirable for females than it was for males and so 'qualified' for the BSRI). In retrospect, it can be seen that the only way for participants to indicate negative evaluations was to rate a positive trait as 'a very poor description'. This is not wholly satisfactory as for items like 'independent' a rating of '1' does not necessarily mean that the participant sees the character as over-dependent. On the attribution items, a negative evaluation could be imposed on the characters by endorsing an unstable, external explanation (e.g. he was lucky) as being 'a very likely explanation'. However, it is not clear that everybody sees unstable and external explanations as being undesirable. As mentioned previously, the way in which people view the explanations on each dimension may be different anyway. Participants may even use their own dimensions to categorise explanations. Despite this problem, it is clear that the predicted differences in attitudes towards the male and female characters were not elicited by this questionnaire.

An additional problem with this questionnaire was that, in asking the participants to describe the character on a list of traits, it was not possible to ascertain how the participants actually felt about the character in any emotional or evaluative respect. One of the main predictions arising from research into females and sport is that one of the barriers to females taking part in sport is a fairly widely held *negative* stereotype. Most attitude theories acknowledge that there are at least two major components to attitudes - a cognitive component and an affective or evaluative component. It may

be that participants accept that dominant and competitive women take part in sport but they may have positive or negative feelings about this. This questionnaire was not able to test participants' affective evaluations of the characters.

The explanations given above for the results in the present study are mainly concerned with specific issues of the questionnaire design. It is possible that there are more general problems with the methodology employed in this study, and it is because of these that the children demonstrated no evidence of gender stereotyping. For instance, Baron and Byrne (1994:p141) suggest that

...the presence of a hot bubbling pizza may activate your attitude towards pizza in a powerful fashion; coming across a newspaper article about the budget deficit may activate your attitude toward the deficit to a much weaker degree.

It may be the case that this questionnaire was not 'meaningful' enough to the participants, and was not powerful in eliciting gender stereotypes. Possibly, participants would show stereotypic attitudes towards females in sport if they were actually watching a sporting event. It is also possible that an interview would be more effective in eliciting gender stereotypes, by asking participants questions which are more 'meaningful' to them individually. Nevertheless, the present study was successful in eliciting stereotypes regarding different hobbies. At some level therefore, it can be assumed that the participants were able to generate a meaningful impression of the characters in the vignettes.

Studies derived from Milgram's lost letter technique (Milgram et al. 1965) suggest that disguised attitude measures are only able to investigate attitudes when they are

very strongly held. It is possible that children do hold stereotypes of males and females in sport, but that their attitudes are not as strong or as negative as predicted. It may be that the method employed by the present study is only able to elicit strong negative attitudes.

Yzerbyt et al. (1994), the authors of social judgeability theory, suggest that people only make judgements about other people when they feel that they are in a position to do so (see Chapter One). Part of feeling entitled to judge is having enough information about the target individual to make an adequate impression. It may be that the children in this study did not feel they had enough information about the characters to make the judgements that they were asked to. Yzerbyt et al. propose that, in making judgements, we follow social norms, one of which states that 'it is wrong to make judgements about other people when you only have categorical information about them'. Perhaps the children did not have to know the nature of the study to know that, in general, it is wrong to judge people based on their category membership (i.e. whether they are male or female). This norm however, may not have extended to stopping the children from making judgements about people who do various different hobbies. Suggesting that a swimmer is more athletic and competitive than a violinist is unlikely to violate any social taboos.

Summary

There are a variety of explanations for the results of Study One. It may be that, contrary to expectations, children do not hold stereotypic views of males and females in sport. An argument could be made that education and social changes have resulted in the stereotypes of yesterday being eliminated amongst today's generation of schoolchildren. However, much further work must be carried out before such a conclusion can be drawn. For instance, there are various aspects of the questionnaire

and the methodology employed that could explain why no gender stereotypes were found and these possibilities should be explored. Firstly, disguised attitude measures may only be able to elicit strongly held attitudes. Children's attitudes towards males and females in sport may not be as strong as they were thought to be. Secondly, the vignette investigated attitudes towards females taking part in a neutral sport, namely swimming. Negative stereotypes may only be held concerning females in sports usually associated with males. Thirdly, the items in the questionnaire did not allow for the affective, emotional aspect of attitudes to be investigated directly. How people *feel* about females in sport is as important as what people *believe* about females in sport. Fourthly, the information in the vignette, while intended to be neutral and 'non-diagnostic' may have given a strong impression of characters that were hard working and studious. This impression may have over-shadowed any effects of varying the sex of the character and the hobby of the character. Fifthly, the children may have withheld their views in order to conform to social norms which suggest that it is wrong to judge people when you do not have much information about them. Finally, it cannot be discounted that the children involved in this study guessed or were told be their friends the purpose of the research, and this resulted in them providing socially desirable answers. While feedback from the children suggests that this is not the case, it is not possible to tell if they were providing 'socially desirable feedback'. Study Two aimed to test some of these possible explanations as to why this questionnaire did not elicit gender stereotypes of males and females in sport.

Conclusions

Predictions that the female character who was described as a sport participant would be negatively evaluated or stereotyped (compared to the male sport participant, the female musician and the characters described as doing no serious hobby) were not supported. Sex of participant and age of participant had no consistent significant effects on the evaluations of the characters. Evidence that the children evaluated the characters in line with stereotypes of hobbies suggests that the children did understand

the questionnaire and were able to form impressions of the characters and subsequently describe these characters. However, from these results it was not possible to conclude that children do not hold stereotypic views of male and female sports participants. There are issues regarding the design of the questionnaire and the use of disguised attitude measures which remain unresolved and these were explored in the later studies.

CHAPTER FOUR: THE GOLDBERG PARADIGM (MK II): REBECCA AND DANIEL PLAY RUGBY AND GO SWIMMING

Introduction to Study Two

Given that gender stereotypes have been widely described as 'barriers' to female participation in sport, it is important to explain why Study One did not find any evidence of gender stereotyping. The two aims of Study Two were, therefore, intertwined. One aim was to investigate the stereotypes that children hold of males and females in sport, and this was an end in itself. However, contingent on eliciting gender stereotypes, Study Two also aimed to provide explanations as to why Study One found no evidence of gender stereotyping. It was hoped that Study Two could provide valuable information regarding the use of the Goldberg Paradigm as a method for assessing stereotypes.

It is important to note that the possibility that the participants in Study One simply did not hold stereotypic views of males and females in sport cannot be eliminated. While this would contradict predictions made from a large body of research and the views of policy making bodies (see Chapter One), it is possible that children today hold very different views of the sexes and that stereotypes are either much weaker than previously or have been eliminated completely. However unlikely this may seem, this is a possible explanation for the findings from Study One.

Study Two, therefore, utilised the same methodology as Study One, but the questionnaire was considerably transformed in order to investigate some of the possible explanations of why Study One found no evidence of gender stereotyping.

One of the possible problems with the vignette in Study One was that the information regarding the characters' sporting involvement may have been overshadowed by the other information, resulting in children generating impressions of the characters as being 'hard working' and 'studious' rather than as 'sporting' or 'musical'. In terms of Fiske and Neuberg's model of impression formation (Fiske and Neuberg, 1990), sport was less salient than other features of the vignette and, therefore, stereotypes reflecting the interaction of gender with sport may not have been evoked. In Study Two, information about the character's sport involvement featured prominently. This was to leave the participants in no doubt that the characters that they were reading about were seriously involved in sport. It was intended that the chances of the character being categorised in other terms would be greatly reduced.

A further feature of the vignette from Study One which was changed for Study Two was the sport that the characters were described as participating in. In Study One, a 'neutral' sport was chosen to assess whether the female character would be denigrated or negatively stereotyped for taking part in sport *per se*. However, it may be that females will only be denigrated when taking part in sports that are usually associated with males (Brawley, Landers, Miller and Kearns, 1979). For this reason, Study Two investigated the different evaluations of male and female characters who were described as taking part in a 'neutral' sport and a 'masculine' sport (i.e. a sport usually associated with males more than females).

A further development in the questionnaire used in Study Two, was the introduction of items intended to gauge an emotional or affective reaction towards the characters

in the vignettes. In Study One it may have been very difficult for the participants to evaluate the characters in negative ways. The majority of the traits (because they were derived from the Bem Sex Role Inventory) tended to have positive connotations. In Study Two, additional trait items were added to those from Study One which could be interpreted as having negative connotations e.g. 'strange', 'obsessed' and 'unusual'. Attributional items were also included again, and a further section of 'personal' items was added. These items asked straightforward 'emotional' questions like 'How much do you like the character in the story?' and 'Would you like him/her as a friend?'. It was hoped that, by including a greater variety of items, the children, if they wanted to, would be able to evaluate the characters in a negative way. It is not clear whether Study One enabled participants, had they wished, to describe the characters in a directly negative way.

A general problem with the Goldberg paradigm comes from changing the name of the character to manipulate the sex of the character. Names do not just indicate a person's sex. Names also have very strong associations with race, age, intellectual competence and social class, and some names are perceived to be more attractive than others (Kasof, 1993). In Study One, it may have been the case that 'Susan' was perceived to be a much more attractive name than 'Steven', and this may have resulted in more positive evaluations of the female characters. Similarly, 'Steven' may be seen as more old-fashioned than 'Susan' and the male characters may have been evaluated in line with age stereotypes rather than gender stereotypes. Kasof (1993) provides a selection of male and female names matched for attractiveness and intellectual competence. These names, however, were rated by undergraduates in North America

and this severely restricts the generalisability to schoolchildren in North East England. For Study Two it was decided to construct the vignettes using the most popular names for babies in the United Kingdom - Rebecca and Daniel (The Guardian, August 5th 1995). By being matched in popularity at a particular point in time it would then seem quite likely that both names were similar in attractiveness and age connotations. By choosing the most popular names, there is less chance that definitive social class information is also provided. That is, the rarer the name the more likely it would be to associate that name with a particular section of society. Nevertheless, it is important to note that regardless of how well names may appear to be matched for attractiveness, age connotations, social class and race associations, it is unlikely that a male name and a female name are going to convey exactly the same information.

In comparison to Study One, which investigated the attitudes of three age groups (11 and 12-year-olds, 13 and 14-year-olds, 15 and 16-year-olds), Study Two only investigated the attitudes of one age group. Children from Year 9 (13 and 14-year-olds) were selected as this age is particularly important for gender awareness (Durkin, 1995). Additionally, selecting only one age group simplified the design of the study by removing one independent variable, and also reduced the likelihood of encountering literacy problems amongst young children who did not understand the questionnaire. Furthermore, by using participants from one year group it was possible to specify the age of the characters in the vignettes. This was important because feedback from participants in Study One indicated that the schoolchildren encountered some difficulties in forming impressions of the characters in the vignettes when no mention was made of each character's age. This had been deliberately left unspecified

because the participants were in different age groups, and had the character's age been included this may have resulted in the character appearing more similar to the participants in one age group than the others. This in turn may have influenced the evaluations of the characters. For instance, research has found that individuals are attracted to the people that they perceive as similar to themselves (Hogg and Vaughan, 1995). By using only one age group in Study Two it was possible to include the character's age in the vignette without causing confounding interactions.

Specifying the characters' ages was important as there is a possibility that counterstereotypic behaviour amongst girls may be tolerated more when they are younger than when they are older. Young girls who enjoy active pursuits more commonly associated with males may be labelled as 'tomboys', though it should be noted that such a label is not necessarily a negative evaluation. However, as girls pass through puberty and into adolescence there are strong expectations that they should behave in more feminine ways, and that they should lose the 'tomboy' image. In Study One, it is possible that the participants could dismiss the female character's sport involvement as a 'young girl's phase'. For this reason, the characters in the vignette in Study Two were described as being 15 years old. In this way, the female characters are described as being serious sports participants, and at age 15 are less likely to be dismissed as 'tomboys'.

In summary, Study Two utilised the same methodology as Study One, but developed the questionnaire in three distinct ways. Firstly, the vignette described a character and their sport involvement without any other extraneous information. Secondly, the male

and female characters were described as playing either a masculine sport or a neutral sport, and thirdly, the rating scales were extended to include a wider variety of positive, negative, emotional and evaluative items.

Method

Design

This study comprised a 2 x 2 x 2 design varying the sex of the character in the story (Daniel Smith or Rebecca Smith), the sport of the character in the story (rugby or swimming) and the sex of the participant (male or female). Participants were tested in groups of around 20 with all four questionnaires distributed randomly within each group. Each group contained both male and female participants.

Questionnaire Design

In total there were four different questionnaires reflecting the four possible combinations of character sex and sports (male rugby player, female rugby player, male swimmer, female swimmer). The items in each questionnaire were exactly the same except for the changes required by the factors being manipulated. An example of a questionnaire used in Study Two can be found in Appendix B.

Vignette Design

As mentioned in the introduction, the vignettes were required to portray the male and female characters as being seriously committed to sport. The vignettes, therefore, described how the characters were selected to go to a summer school for their particular sport. As well as describing the sort of activities the characters would be doing, the vignettes included quotes from the character about their future ambitions and from the coaches regarding their feelings towards the summer school.

A pilot study was carried out with a group of 81 undergraduate students to establish the two sports to be included in the vignettes. From a list of 40 common sports, rugby and swimming were selected. Apart from boxing, rugby was the sport that was rated as the most 'masculine', i.e. more appropriate for males than females. Boxing was not chosen for this study because of the many controversies surrounding this sport. Many people believe that boxing is an inappropriate sport for anybody, regardless of their sex. Had boxing been included in the vignette, it would have been difficult to distinguish between general disapproval and disapproval based on gender stereotyping. Swimming was chosen as the sport which was rated as being the most neutral, i.e. appropriate for both males and females.

As in Study One, the information presented in the vignettes was intended to enable the participants to build an impression of the character but not to give any diagnostic information which could directly answer the questions following the vignette. That is, the information provided was such that participants would need to make inferences about the characters to be able to complete the questionnaire

The complete vignette was as follows:

Rebecca/Daniel Smith, aged 15, a pupil at West Holden Comprehensive School, has just been selected to take part in a summer school organised by Cambridgeshire Schools Swimming/Rugby Association. Along with 45 other girls/boys from around the county, Rebecca/Daniel will spend three weeks at Foresham Hall. Each day the swimmers/rugby players will practise the skills and techniques essential to swimming/rugby and follow a challenging fitness training programme. Rebecca/Daniel said "I am really looking forward to it. Three weeks of hard training will be tough, but I am sure I will learn a lot from it". Rebecca/Daniel was unsure how far he/she would go in the sport. "Well it's a bit early to say. I hope to swim/play for the county team one day. Some of the girls/boys at the summer school will be very good, and I'll need to work hard to keep up with them". The coaches who organised the summer school see it as an important way of improving standards throughout the sport. "It's great to get them all together" said one coach. "That way they all learn from each other, and because they are so competitive they push each other to do their best. It can only be good for the sport of swimming/rugby."

Item Design

A possible shortcoming of the first study was the lack of items designed to elicit the evaluative components of the participants' attitudes towards the sexes in sport. Consequently, the second questionnaire included more questions of greater variety than in the first questionnaire, split into 6 main sections. A summary of the structure of the questionnaire is shown in Table 4.1.

Vignette		Examples
Section 1	Trait Items	Aggressive, Independent, Gentle
Section 2	Attributional Items	Why do you think Daniel/Rebecca was selected for the summer school?
Section 3	Attributional Items	Why do you think Daniel/Rebecca wanted to go to the summer school?
Section 4	Evaluative Items	How much do you like Daniel/Rebecca?
Section 5	Attributional Items	Why does Rebecca/Daniel go swimming/play rugby?
Section 6	Personal details	

Table 4.1 Questionnaire Structure

The first section of questions comprised a list of traits, and as in Study One the participants were asked to rate these on a five point Likert-type scale ranging from *a very poor description* to *a very good description*. There were 19 adjectives in all, including words such as 'strange', 'obsessed' and 'unusual', which could be construed as having negative connotations, as well as 'feminine' and 'masculine'. Table 4.2 shows the trait items along with possible classifications in terms of whether the trait is perceived as being a 'masculine' or 'feminine' trait (from BSRI classification - see Chapter 3). The 'desirability' of each trait is also shown - desirable traits are those that are seen to be positive, undesirable traits are those that are seen to be negative. Study One asked participants to use almost entirely positive traits to describe the characters, whereas Study Two was designed to give participants the opportunity to use both positive and negative traits. From this study it was not intended to establish whether the participants perceived the traits in line with the classifications presented in Table 4.2. However, the important feature of this study is whether children used the traits to differentiate between the different characters, rather than the ways in which they perceived individual statements. These classifications are not intended to be

definitive but are presented to demonstrate the variety of traits presented to the participants in this study.

	Desirability (positive or negative)	Masculine or Feminine
Independent	Positive	Masculine
Confident	Positive	Masculine
Competitive	Positive	Masculine
Strange	Negative	Neutral
Attractive	Positive	Neutral
Masculine	Positive for males Negative for females	Masculine
Gentle	Positive	Feminine
Honest	Positive	Neutral
Friendly	Positive	Neutral
Selfish	Negative	Neutral
Feminine	Positive for females Negative for males	Feminine
Bigheaded	Negative	Neutral
Hard working	Positive	Neutral
Obsessed	Negative	Neutral
Aggressive	Positive	Masculine
Daring	Positive	Masculine
Unusual	Negative	Neutral
Athletic	Positive	Masculine
Trendy	Positive	Neutral

Table 4.2 Trait Items with Possible Classification on Desirability and Masculine/Feminine Dimensions

Attribution Items

Sections 2, 3 and 5 comprised further attributional items to allow participants to indicate their perceptions of why the character was selected for the summer school,

why the character wanted to go to the summer school, and why the character took part in the sport. As in Study One, these possible explanations were constructed around the dimensions of locus of control (internal/external) and stability over time (stable/unstable), and again participants rated these explanations on five point Likert-type scales ranging from *a very unlikely explanation* to *a very likely explanation*.

Section 2 of the questionnaire asked the participants 'Why do you think Rebecca/Daniel was selected for the summer school?' and asked them to rate the likelihood of possible explanations. Table 4.3 shows these 'Attribution Items' and a possible classification for each item in terms of its 'desirability' as an explanation. For example, 'There was a place for anybody who wanted to go' was designed to be an undesirable or negative attribution to explain why the character was selected for the summer school. On the other hand, 'He was the best swimmer in his school' was intended to be a desirable or positive explanation.

As with the trait items, it was not intended to establish whether the participants perceived the attributions in the same ways as was intended in their design. Russell (1982) suggests that some researchers have assumed that they can predict how participants perceive causes, and in doing so have committed what he calls the 'fundamental attribution researcher error'. While in the design of the questionnaire, items with an external locus of control were intended to be negative explanations, it is quite possible that some of the participants perceived the causes in other ways. What was important for this study, was not the ways in which individual items were interpreted by the participants, but whether the participants used the items to evaluate the different characters in different ways. The intention was, therefore, to provide the

participants with as many different kinds of statements, both positive and negative, with which to describe the characters. This caveat remains for all the classifications proposed for the items in Study Two.

	Locus of Control (internal or external)	Stability (stable or unstable)	Desirability (positive or negative)
He was the best swimmer in his school.	internal	stable	positive
His parents wanted him to go.	external	unstable	negative
His parents are good friends with the summer school coaches.	external	stable	negative
There was a place for anybody who wanted to go.	external	unstable	negative
No other boy in his school wanted to go.	external	unstable	negative
He was lucky.	external	unstable	negative
He is a natural all-round sports player.	internal	stable	positive
He tries hard at all sports.	internal	stable	positive
He was in the right place at the right time.	external	unstable	negative

Note : For 'masculine' sport conditions 'swimmer' and 'swimming' were replaced with 'rugby player' and 'playing rugby', and for conditions with female characters the name and all pronouns were changed.

Table 4.3 Attribution Items (Section 2) and Possible Classification on Stability, Locus of Control and Desirability Dimensions

Section 3 asked participants 'Why do you think Rebecca/Daniel wanted to go to the summer school?'. They were then asked to rate the likelihood of some possible explanations which are listed in Table 4.4 along with possible ways in which these explanations could be classified. These items were rated on five point Likert-type scales ranging from 1, a *very unlikely explanation* to 5, a *very likely explanation*.

	Locus of Control (internal or external)	Stability (stable or unstable)	Desirability (positive or negative)
He enjoys physical exercise	Internal	Stable	Positive
He wants to become a better swimmer	Internal	Unstable	Positive
It's one step towards his ambition of swimming for the county team	Internal	Unstable	Positive
His parents encouraged him to go.	External	Unstable	Negative
His older brother has been to a summer school before.	External	Unstable	Negative
He wanted a holiday in the country.	External	Unstable	Negative
His father was a swimmer and he wants to follow in his footsteps.	External	Unstable	Positive
He thinks that being sporty will make him popular at school.	External	Unstable	Negative
He has some friends who are also going to the summer school.	External	Unstable	Negative
He hopes to make some new friends who also swim.	External	Unstable	Positive

Note : For 'masculine' sport conditions 'swimmer' and 'swimming' were replaced with 'rugby player' and 'playing rugby', and for conditions with female characters the name and all pronouns were changed.

Table 4.4 Attribution Items (Section 3) and Possible Classification on Stability, Locus of Control and Desirability Dimensions

Section 5 asked participants to think of some reasons why Rebecca/Daniel went swimming/played rugby. They were then asked to rate the likelihood of some possible explanations which are listed in Table 4.5 along with possible ways in which these explanations could be classified. These items were rated on five point Likert-type scales ranging from 1, a *very unlikely explanation* to 5, a *very likely explanation*.

	Locus of Control (internal or external)	Stability (stable or unstable)	Desirability (positive or negative)
He does it to keep fit.	Internal	Unstable	positive
He enjoys competitions.	Internal	Stable	positive
He enjoys physical exercise.	Internal	Stable	positive
The training improves his good looks.	Internal	Unstable	negative
Being sporty means he is popular at school.	External	Unstable	negative
A lot of his friends also go swimming so he wants to join in too.	External	Unstable	negative

Note : For 'masculine' sport conditions 'swimmer' and 'swimming' were replaced with 'rugby player' and 'playing rugby', and for conditions with female characters the name and all pronouns were changed.

Table 4.5 Attributional Items (Section 5) and Possible Classification on Stability, Locus of Control and Desirability Dimensions.

Section 4 included evaluative and projective items designed to find out how participants felt towards the characters. Amongst these items were questions asking how much the participant liked the character, how well they thought the character would do in the future on exams, and how likely it was that the character would fulfil

their ambitions. All but one of these questions were answered using five point Likert-type scales. The exception was the question asking how many hours the character should spend each week on their sport. The complete set of questions along with the labels for the extreme points of the scales are shown in Table 4.6.

The sixth and final section of the questionnaire asked the children personal details such as their age and sex, and also how active in sport they perceived themselves to be. The questionnaire in Study One also attempted to gauge the seriousness of the participants' own sport participation but there were comprehension problems with the questions. Section 6 of the questionnaire in Study Two was designed to overcome these problems by giving examples of possible answers, and by specifying that participation in seasonal sports should be reported even if the participants were not currently participating. They were asked to rate their own level of activity in sport on a five point Likert-type scale ranging from 1, (*not at all active in sport*) to 5, (*very active in sport*). They were also asked to list the three sports which they played most often and to indicate how often they played each ('At least once a week', 'At least once a fortnight' or 'At least once a month'). For the sport they played most often, they were asked to indicate the 'level' at which they participated. This was an open ended question but they were given examples to prompt them ('e.g. play with friends, play for school, play for county team, etc.'). The aim of these questions was to categorise the children as serious or casual sports participants as this may have influenced the way in which they completed the questionnaire.

Instructions To Participants

For the trait items the written instructions to participants were identical to those in Study One (see Chapter 3). As previously, the five point scale was explained and participants were informed that if they did not understand a word they should mark it with a cross and then move on to the next question.

For the attributional items the instructions were identical to those used in Study One except that 'In the story Steven/Susan won a prize for his geography project. Can you think of some reasons why he won the prize?' was replaced with 'Why do you think Daniel/Rebecca was selected for the summer school?', 'Why do you think Daniel/Rebecca wanted to go to the summer school?' and 'Can you think of some reasons why Daniel/Rebecca goes swimming/plays rugby?' in the corresponding sections. Participants were asked to rate the likelihood of each explanation on five point Likert-type scales ranging from *very unlikely explanation* to *very likely explanation*.

For the evaluative questions (Section 4) the instructions were as follows:

In this section I want you to try and imagine what kind of person Daniel/Rebecca is and then answer the following questions...

At the end of the questionnaire participants were thanked for their help and were given the opportunity to write any comments they had about the story or the questions.

Participants

This questionnaire was presented to 153 children in Year 9 (mean age 13.5 years, SD 0.58) of a comprehensive school in the North East of England.

Pilot Study

A pilot study was carried out with 21 Year 9 (13 and 14 year old) children in a different comprehensive school from the one used for the main data collection. The pilot study was successful and no changes were made to the questionnaire, and it was not anticipated that there would be any problems testing the main sample at this first school. However, after two sessions the researcher, after consultation with school staff, decided to withdraw and to approach another school. The main problems encountered were that the children did not concentrate on their own questionnaire and they discussed their answers. This meant that they became aware of the different vignettes and the purpose of the study was exposed. In the second school it was possible to ensure that the children did work alone and complete the questionnaire without discussion. The children in the original school, however, did not demonstrate any problems in understanding the questionnaire and so the vignette and the items remained unchanged for the main sample.

Procedure

The procedure for Study Two was identical to that utilised in Study One (see Chapter Three).

Results

The results of this second study echo those of the first. While there is evidence to suggest that the children did understand the questionnaire and that the manipulations were effective, there is no evidence to suggest that the female character in either rugby or swimming was denigrated or negatively evaluated in any way. Table 4.7 shows the mean scores for the trait items for each of the four characters (shaded boxes indicate results that are discussed in the text). These mean scores demonstrate that the characters were, in fact, described in very similar ways. While there were some main effects for sex of character and hobby of character, there were no consistent interaction effects of these two factors which would suggest that the female rugby player was being evaluated differently from the other characters.

	Male character		Female character	
	swimming	rugby	swimming	rugby
independent	3.31 (0.86) n=36	3.38 (0.85) n=39	3.29 (0.80) n=34	3.46 (0.77) n=37
confident	4.37 (0.67) n=38	4.03 (0.96) n=39	4.08 (0.89) n=37	3.95 (0.98) n=38
competitive	4.36 (0.90) n=36	4.29 (0.84) n=38	4.06 (0.97) n=35	4.18 (0.87) n=38
strange	1.55 (0.89) n=38	1.67 (0.90) n=39	1.61 (0.84) n=36	1.95 (0.98) n=38
attractive	2.24 (1.28) n=37	2.29 (1.18) n=38	2.56 (1.13) n=36	2.39 (1.05) n=38
masculine	2.78 (1.20) n=36	3.15 (1.19) n=40	2.12 (1.17) n=34	2.94 (1.30) n=34

Cell contents: Mean, standard deviation, number of participants
Rating scale: 1=very poor description to 5=very good description

Table 4.7 Mean scores for Trait Items by Sex of Character and Sport of Character

	Male character		Female character	
	swimming	rugby	swimming	rugby
gentle	2.68 (1.25) n=37	2.28 (1.11) n=40	3.14 (1.06) n=37	2.50 (1.11) n=38
honest	3.92 (1.16) n=37	3.92 (0.87) n=39	3.97 (1.12) n=35	3.94 (1.07) n=36
friendly	4.05 (1.10) n=37	3.65 (1.19) n=40	4.24 (0.72) n=37	4.03 (0.83) n=37
selfish	1.66 (0.99) n=38	1.93 (0.97) n=40	1.73 (1.10) n=37	1.42 (0.55) n=38
feminine	1.59 (1.05) n=34	1.67 (0.87) n=39	3.36 (1.20) n=36	2.46 (1.32) n=37
bigheaded	1.92 (1.14) n=37	2.11 (0.92) n=38	1.92 (1.04) n=37	1.82 (0.98) n=38
hard working	4.46 (0.80) n=37	4.51 (0.76) n=39	4.65 (0.54) n=37	4.59 (0.69) n=37
obsessed	2.58 (1.25) n=36	2.95 (1.21) n=39	2.71 (1.18) n=35	2.66 (1.19) n=38
aggressive	1.84 (0.86) n=38	2.45 (1.18) n=40	1.69 (0.79) n=36	2.66 (1.21) n=38
daring	2.65 (1.03) n=37	3.10 (1.10) n=40	2.74 (1.09) n=35	3.42 (1.08) n=38
unusual	2.03 (1.03) n=36	1.82 (0.94) n=39	1.83 (0.86) n=35	2.38 (1.28) n=37
athletic	4.61 (0.64) n=36	4.56 (0.72) n=39	4.62 (0.55) n=37	4.62 (0.64) n=37
trendy	2.47 (1.08) n=38	2.53 (1.09) n=40	2.86 (1.06) n=37	2.39 (0.92) n=38

Cell contents: Mean, standard deviation, number of participants

Rating scale: 1=very poor description to 5=very good description

Table 4.7 (Contd.) Mean scores for Trait Items by Sex of Character and Sport of Character

Table 4.8 shows the mean scores for the four different vignettes for the attribution items from Section 2 of the questionnaire (Why do you think Daniel/Rebecca was selected for the summer School?). Again, there were no consistent results from the attributional items to suggest that the female characters were denigrated in comparison to the male characters or that the female rugby player was denigrated in comparison to the female swimmer.

	Male character		Female character	
	swimming	rugby	swimming	rugby
She/he was the best swimmer/rugby player in her/his school	3.95 (0.98) n=38	3.60 (1.26) n=40	3.59 (1.26) n=37	3.63 (0.91) n=38
Her/his parents wanted her/him to go	2.66 (1.21) n=38	2.25 (1.06) n=40	2.51 (1.19) n=37	2.58 (0.98) n=38
Her/his parents are good friends with the summer school coaches	2.13 (1.26) n=38	2.13 (1.20) n=39	2.11 (1.29) n=37	1.97 (0.91) n=38
There was a place for anybody who wanted to go	1.95 (1.11) n=38	2.36 (1.50) n=39	2.35 (1.36) n=37	2.47 (1.45) n=38
No other girl/boy in her/his school wanted to go	1.71 (1.04) n=38	1.83 (1.01) n=40	1.84 (0.90) n=37	1.97 (1.05) n=38
She/he was lucky	2.66 (1.26) n=38	3.08 (1.31) n=40	2.95 (1.10) n=37	2.37 (0.97) n=38
She/he is a natural all-round sports player	3.54 (0.99) n=37	3.33 (1.10) n=40	3.73 (0.90) n=37	3.58 (1.00) n=38
She/he tries hard at all sports	4.14 (0.95) n=37	4.20 (0.91) n=40	4.46 (0.61) n=37	4.08 (0.75) n=38
She/he was in the right place at the right time	2.45 (1.13) n=38	2.33 (1.21) n=40	2.49 (1.12) n=37	2.32 (0.87) n=38

Cell contents: Mean, standard deviation, number of participants

Rating scale: 1=very unlikely explanation, 5=very likely explanation

Table 4.8 Mean Scores for Attribution Items (Why do you think Rebecca/Daniel was selected for the summer school) by Sex of Character and Sport of Character.

Table 4.9 shows the mean scores for the four different vignettes for the attribution items from Section 3 of the questionnaire (Why do you think Daniel/Rebecca wanted to go to the summer School?). As with the previous two sections these results also demonstrate that all four characters were described in very similar ways.

	Male character		Female character	
	swimming	rugby	swimming	rugby
She/he enjoys physical exercise	4.16 (0.97) n=38	4.08 (0.92) n=40	4.00 (0.91) n=37	4.11 (0.80) n=38
She/he wants to become a better swimmer/rugby player	4.79 (0.41) n=38	4.55 (0.71) n=40	4.65 (0.59) n=37	4.58 (0.55) n=38
It's one step towards her/his ambition of playing/swimming for the county team	4.63 (0.71) n=38	4.58 (0.59) n=40	4.54 (0.65) n=37	4.68 (0.53) n=38
Her/his parents encouraged her/him to go	3.08 (1.17) n=38	2.65 (0.83) n=40	3.08 (1.01) n=37	2.89 (0.95) n=38
Her/his older brother has been to a summer school before	2.50 (1.22) n=38	2.30 (0.94) n=40	2.65 (1.16) n=37	2.39 (0.92) n=38
She/he wanted a holiday in the country	2.08 (1.38) n=37	1.65 (1.00) n=40	2.22 (1.18) n=37	2.00 (.99) n=38
Her/his father is a rugby player/swimmer and she/he wants to follow in his footsteps	2.32 (1.07) n=38	2.38 (1.19) n=40	2.64 (1.10) n=36	2.50 (1.08) n=38
She/he thinks that being sporty will make her/him popular at school	2.37 (1.10) n=38	2.23 (1.14) n=40	2.32 (1.20) n=37	2.24 (1.15) n=38
She/he has some friends who are also going to the summer school	2.79 (1.07) n=38	2.75 (1.15) n=40	3.00 (1.33) n=37	2.68 (1.07) n=38
She hopes to make new friends who also play rugby/go swimming	3.87 (0.88) n=38	3.78 (1.00) n=40	3.95 (0.94) n=37	3.84 (0.95) n=38

Cell contents: Mean, standard deviation, number of participants

Rating scale: 1=very unlikely explanation, 5=very likely explanation

Table 4.9 Mean scores for Attribution Items (Why do you think Rebecca/Daniel wanted to go to the summer school?) by Sex of Character and Sport of Character

	Male character		Female character	
	Swimming	Rugby	Swimming	Rugby
How much do you like Rebecca/Daniel?	3.21 (0.99) n=38	3.33 (0.89) n=40	3.54 (0.93) n=37	3.42 (0.79) n=38
Would you like her/him as a friend?	3.34 (1.17) n=38	3.33 (1.01) n=39	3.57 (1.09) n=37	3.32 (0.90) n=38
How hard do you think she/he works at her/his schoolwork?	3.63 (1.05) n=38	3.68 (0.92) n=40	4.24 (0.76) n=37	3.92 (0.85) n=38
How well do you think she/he will do in her/his exams?	3.37 (0.85) n=38	3.53 (0.72) n=40	3.92 (0.64) n=37	3.55 (0.76) n=38
Do you think she would be a good captain for her team?	4.34 (0.81) n=38	3.98 (1.00) n=40	4.30 (0.85) n=37	4.21 (0.70) n=38
Do you think Rebecca/Daniel should do more work at swimming/rugby or less work at swimming/rugby?	3.55 (0.80) n=38	3.75 (0.84) n=40	3.62 (0.72) n=37	3.39 (0.79) n=38
*How many hours a week do you think Rebecca/Daniel should spend training?	12.16 (15.99) n=37	10.74 (14.74) n=34	17.11 (25.58) n=37	10.79 (12.25) n=38
Do you think Rebecca/Daniel should spend more or less time training?	3.43 (0.80) n=37	3.31 (0.92) n=39	3.38 (0.95) n=37	3.37 (0.71) n=38
Do you think Rebecca/Daniel should spend more or less time doing homework?	3.71 (0.87) n=38	3.40 (1.15) n=40	3.46 (1.02) n=37	3.59 (0.90) n=37
Do you think Rebecca/Daniel should spend more or less time going out with friends?	3.34 (1.12) n=38	3.23 (1.05) n=40	3.16 (1.01) n=37	3.42 (0.98) n=38
Do you think Rebecca/Daniel should spend more or less time watching T.V.?	2.34 (1.07) n=38	2.25 (1.15) n=40	2.16 (1.09) n=37	2.11 (0.92) n=38

Cell contents: Mean, standard deviation, number of participants

Rating scale: 1='negative' evaluation, 5='positive' evaluation

*this item was not answered using a 5 point Likert scale. Participants were asked to write the exact number of hours.

Table 4.10 Mean Scores for Evaluative items by Sex of Character and Sport of Character

	Male character		Female character	
	Swimming	Rugby	Swimming	Rugby
How likely do you think it is that Rebecca/Daniel will play/swim for the county team?	4.14 (0.75) n=37	3.73 (1.06) n=40	4.11 (0.99) n=37	3.87 (0.93) n=38
How likely do you think it is that Rebecca/Daniel will get married?	3.51 (1.02) n=37	3.78 (0.89) n=40	3.47 (1.06) n=36	3.47 (0.98) n=38
How likely do you think it is that Rebecca/Daniel will go to college or university?	3.49 (0.99) n=37	3.28 (1.04) n=40	3.89 (0.99) n=37	3.63 (1.00) n=38
How likely do you think it is that Rebecca/Daniel will have children?	3.56 (0.97) n=36	3.40 (1.06) n=40	3.42 (1.05) n=36	3.34 (0.94) n=38
How likely do you think it is that Rebecca/Daniel will become a well known sports star?	3.30 (0.94) n=37	3.00 (1.30) n=40	3.36 (1.17) n=36	3.11 (1.18) n=38
How likely do you think it is that Rebecca/Daniel will give up her/his sport when she/he leaves school?	1.97 (1.26) n=38	1.85 (0.87) n=39	1.89 (1.01) n=36	2.13 (1.04) n=38
How likely do you think it is that Rebecca/Daniel will represent her/his country at swimming/rugby?	3.59 (1.07) n=37	3.28 (1.43) n=39	3.69 (1.17) n=36	3.03 (1.05) n=38
How likely do you think it is that Rebecca/Daniel will become a swimming/rugby coach?	3.82 (0.73) n=38	3.51 (1.30) n=39	3.72 (1.06) n=36	3.71 (0.87) n=38

Cell contents: Mean, standard deviation, number of participants
 Rating scale: 1='negative' evaluation, 5='positive' evaluation

Table 4.10(Contd.) Mean Scores for Evaluative items by Sex of Character and Sport of Character

Table 4.10 shows the mean scores for the four different vignettes for the evaluative items from Section 4 of the questionnaire. Table 4.11 shows the mean scores for the four different vignettes for the attribution items from Section 5 of the questionnaire.

	Male character		Female character	
	Swimming	Rugby	Swimming	Rugby
She/he does it to keep fit	3.42 (0.83) n=38	3.13 (1.00) n=39	3.22 (1.10) n=36	3.13 (0.88) n=38
She/he enjoys competitions	4.03 (0.68) n=38	4.00 (0.89) n=39	4.28 (0.78) n=36	4.03 (0.75) n=38
She/he enjoys physical exercise	4.11 (0.86) n=38	4.05 (0.89) n=39	4.31 (0.67) n=36	4.00 (0.74) n=38
The training improves her/his good looks	2.16 (1.03) n=38	2.10 (1.07) n=39	2.36 (0.96) n=36	1.92 (1.08) n=38
Being sporty means she/he is popular at school	2.29 (0.98) n=38	2.46 (1.14) n=39	2.09 (0.83) n=34	2.13 (1.14) n=38
A lot of her/his friends also play rugby/go swimming so she/he wants to join in too	2.71 (1.06) n=38	2.67 (1.08) n=39	2.81 (1.01) n=36	2.37 (1.05) n=38

Cell contents: Mean, standard deviation, number of participants
 Rating scale: 1=very unlikely explanation, 5=very likely explanation

Table 4.11 Mean Scores for Attribution Items (Why do you think Rebecca/Daniel plays rugby/goes swimming?) by Sex of Character and Sport of Character

Although included in the statistical analysis, sex of participant was not found to have any consistent effects on the descriptions of the characters and so, for simplicity, the results are shown in Tables 4.7, 4.8, 4.9, 4.10 and 4.11 for all participants together. It was hoped to include in the analysis the level at which the participants took part in sport as it was hypothesised that serious sports participants may describe the characters in different ways from less serious sports participants. However, as in Study One the participants had problems with answering the question regarding the level at which they played sport. Despite the more detailed instructions which were

added to the questionnaire after the problems encountered in Study One, participants involved in seasonal sports were still confused about how to answer the question regarding how often they played sport. It was, therefore, not possible to include the level of sports participation in the analysis.

Anova analysis of the ratings for the trait item 'feminine' showed a significant main effect for sex of character, $F(1,137)=44.766, p<0.001$, and a significant interaction between sex of character and sport of character, $F(1,137)=7.06, p<0.001$. Tukey post-hoc analyses revealed that the female rugby player ($M=2.46, SD=1.32$) was rated as being less feminine than the female swimmer ($M=3.36, SD=1.20$), and that both female characters were rated as being more feminine than the male characters (male swimmer $M=1.59, SD=1.05$; Male rugby player $M=1.67, SD=0.87$). These results are illustrated in Figure 4.1.

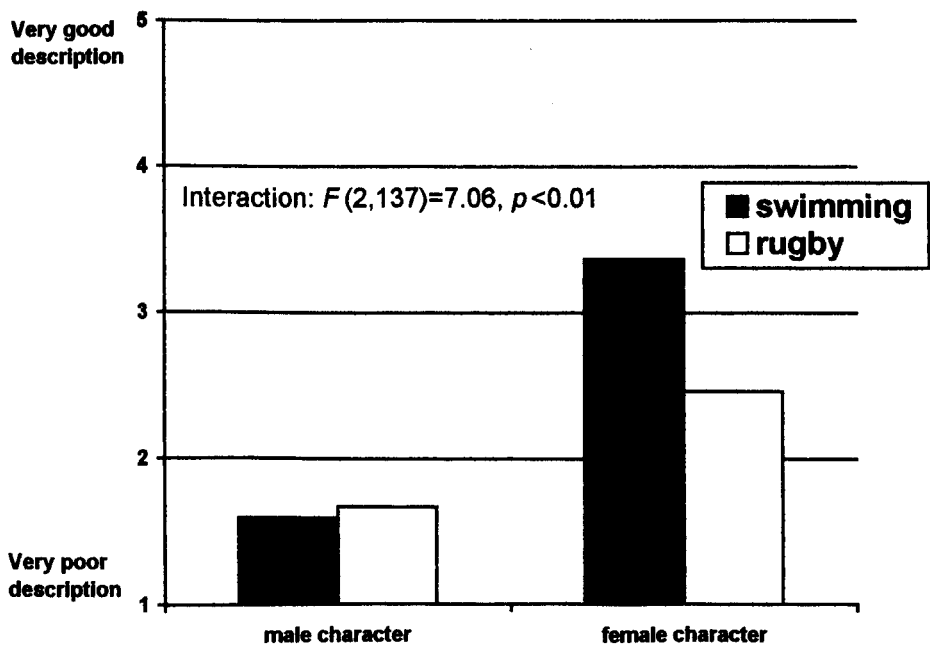


Figure 4.1 Mean Ratings for 'Feminine' by Sex of Character and Sport of Character

It is not possible from this single result to conclude that female rugby players are denigrated because they are seen as 'less feminine'. It is possible that the female rugby player may have been described as being less 'feminine' merely because few females play rugby. This interpretation is backed up by the fact that for the word 'unusual' there was a similar significant interaction for sex of character and sport of character, $F(1,138)=4.693, p<0.05$ (see Figure 4.2). with the female rugby player being rated as the most unusual of the four characters. However, for the word 'strange' there were no significant effects at all. It appears that the children saw the female player as being 'unusual' in the sense of being a rarity rather than in the sense of being odd or weird (which would imply a negative evaluation).

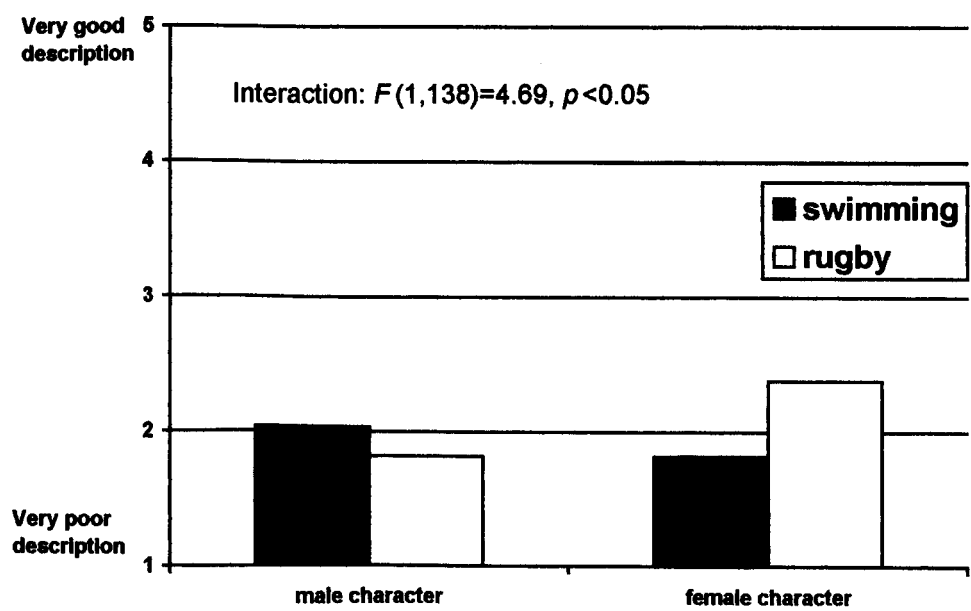


Figure 4.2 Mean ratings for 'Unusual' by Sex of Character and Sport of Character

The results for 'feminine' were not reproduced for 'masculine'. As expected the male rugby player was rated as the most masculine, and the female swimmer was rated the least masculine. However, the female rugby player ($M=2.94$, $SD=1.30$) was actually rated as slightly (but not significantly) more masculine than the male swimmer ($M=2.78$, $SD=1.20$). There was no significant main effect for sex of character, but there was a significant main effect for sport, with the rugby players rated as more masculine than the swimmers, $F(1,143)=8.02$, $p<0.01$.

One interesting interaction between sex of character and sport was found for the explanation 'He/she was lucky' in response to the question 'Why do you think Rebecca/Daniel was selected for the summer school?', $F(1,144)=6.792$, $p<0.05$ (See Figure 4.3). A Tukey post hoc test showed that for the male and female swimmers this explanation was rated similarly. However, for the rugby players, participants rated the male rugby player as luckier than the female rugby player.

One possible explanation for this finding is that participants may have perceived rugby as a highly competitive sport for males and that even talented players would need a little bit of luck to succeed. On the other hand, for female rugby players because so few females play the sport, luck might not be perceived to be as important. However, if this explanation is correct it would be expected that the explanation 'There was a place for anybody who wanted to go' would be seen as more likely for the female rugby player and much less likely for the male rugby player and yet there were no significant main effects or interactions for this item.

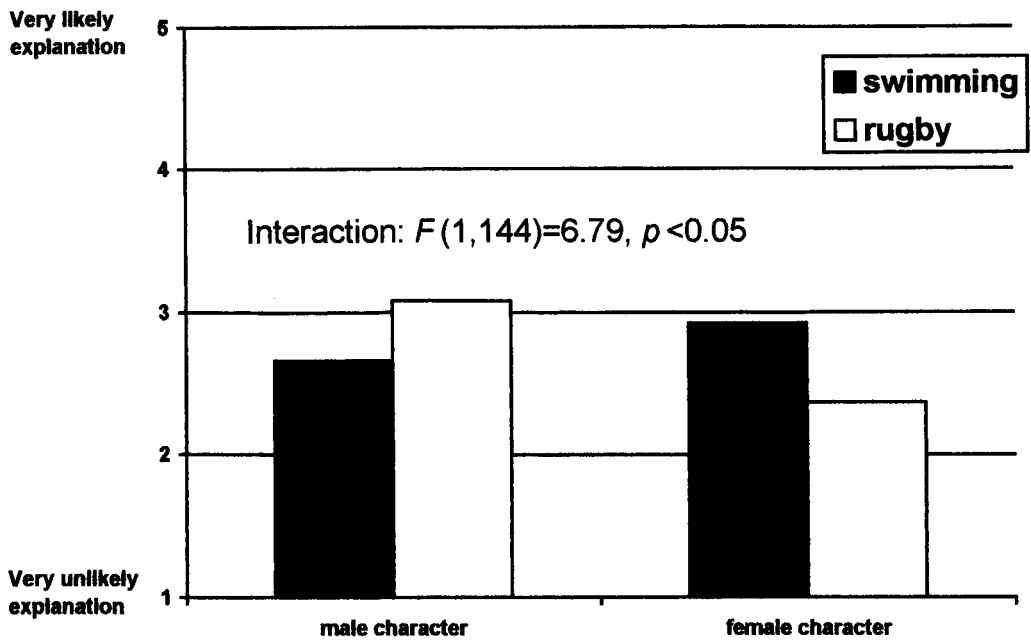


Figure 4.3 Mean Ratings for 'He/She was lucky' by Sex of Character and Sport of Character

There were no other significant interactions between sex of character and sport, though there were some main effects for sport and for sex of character separately. For example, the swimmers were seen as significantly more gentle, $F(1,143)=8.217, p<0.01$, more friendly, $F(1,142)=3.902, p<0.05$, less aggressive, $F(1,143)=24.978, p<0.001$, and less daring, $F(1,142)=10.030, p<0.01$. The female characters were rated as harder workers on the item 'How hard do you think he/she works at his/her schoolwork?', $F(1,144)=8.245, p<0.01$ but there were no significant effects for the trait item 'Hard working'. There was further evidence that the female characters were being evaluated in stereotypic ways as they were rated more likely to go on to college, $F(1,143)=5.487, p<0.05$, and as more likely to do well in their exams, $F(1,144)=5.170, p<0.05$.

Discussion

In summary, the results from Study Two are very similar to those from Study One. There is evidence to suggest that the children did understand the questionnaire in that they demonstrated strong sport stereotypes and some gender stereotypes. These results may be consistent with general views of rugby and swimming, and with figures showing that girls perform better academically at age 16 compared to boys (Office for Standards in Education [OFSTED], 1996). So, while these results tell us little about how females in sport are perceived, this cannot be attributed to the children being unable to complete the questionnaire or not comprehending the task required.

It is clear that while the female rugby player is seen as less feminine and more unusual than a female swimmer there was no evidence to suggest that the female rugby player was viewed in any negative way. This is supported by the fact that the female rugby player was rated as more unusual but not as more strange. Despite using a wide range of traits, both positive and negative, and a large selection of evaluative and projective questions there did not seem to be any generally held negative stereotype which might explain, in part, female under-representation in sport.

From the results of this study, and of Study One, it can be concluded that when completing the questionnaire, the participants did not evaluate the female characters in a negative way compared to the male characters, and also the female taking part in an 'inappropriate' sport was not negatively evaluated compared to a female taking part in an 'appropriate' sport. This still leaves the question of whether children really do not

hold strong stereotypes of males and females in sport or whether children do hold these stereotypes and it is some feature of the questionnaire and its design which is not eliciting differential evaluations. There is evidence to suggest that the manipulations in both Study One and Study Two were effective in that they were able to elicit clear stereotypes regarding the different pastimes and the different sports. The fact that the participants held strong stereotypes of swimmers and rugby players is not entirely surprising. However, the problem still remains that if children do hold stereotypes of the sexes in sport but they are not as strong as predicted then, as Milgram et al. (1965) suggest, a disguised attitude method such as the questionnaire employed may not be effective in eliciting these stereotypes.

Study Two was designed to remedy some of the possible problems with the questionnaire in Study One. For instance, the lack of evidence of gender stereotyping in Study One may have been due to sport not being emphasised in the vignette. Therefore, the vignettes in Study Two were solely about sport and were designed to make the categories 'female rugby player', 'female swimmer', etc., as 'salient' to the readers as possible. It is unlikely, therefore, that the results from Study Two can be explained by suggesting that the participants were unable to form a strong impression of the characters, as it was only possible to categorise the characters in terms of their sport involvement.

A further problem from Study One which was remedied by Study Two was the lack of 'emotional' or 'affective' items to describe the characters. As discussed in Chapter One attitudes are commonly held to consist of three components: the cognitive

component, the affective component and the behavioural component. Whereas Study One concentrated mostly on the cognitive component and the beliefs that the children held about sports participants, Study Two included many items intended to gauge the emotional or affective reaction of the participants towards the characters. Despite this, no evidence of negative attitudes towards the female swimmer or female rugby player were found.

There appear to be two main explanations which could explain the results from Studies One and Two. The first of these is that children do hold negative stereotypes of females in sport and there was some feature of the methodology employed which resulted in the apparently egalitarian attitudes which were reported. However, the results show quite clearly that the methodology was effective in eliciting gender stereotypes and stereotypes of rugby players and swimmers, even though there were no consistent interactions between sex and sport. The children also reported few problems with the methodology. While a few children did not understand some of the items, the vast majority were able to complete the task straightforwardly. It seems, therefore, difficult to attribute the results of Studies One and Two to the design of the questionnaire. However, as mentioned in Chapter Three, there remains the possibility that while the questionnaire provided children with the opportunity to negatively evaluate female sports players they may have 'chosen' not to. That is, according to social judgeability theory (Yzerbyt et al., 1994) the children may be aware of the various social norms regarding when it is appropriate or not to make judgements about other people. The children in this study may be aware of prejudice and discrimination against females and either consciously or unconsciously withheld their

negative feelings towards the female sports players. This possibility is explored in Chapter Six which describes the fourth study in the present research. This study utilised semi-structured interviews to investigate, amongst other issues, children's awareness of norms. Study Five, described in Chapter Seven, utilised a questionnaire to test some of the specific predictions made by social judgeability theory.

The second main possible explanation for the results from Studies One and Two is that attitudes towards females in sport have changed since the research reported in Chapter Three was carried out. While studies such as those by Weinberg et al. (1984) and Parkhouse and Williams (1986) successfully used the Goldberg paradigm to investigate gender stereotyping in sport, this technique was unsuccessful in the present study. It may be that the negative attitudes towards females in sport which were highlighted by the Sports Council (1992) as 'barriers' to participation have either weakened considerably or have disappeared altogether. It is possible that previous generations held strong negative attitudes towards females in sport and these contributed to the participation patterns described in Study One. However since these patterns have been highlighted, the education of subsequent generations may have been successful in eliminating (to a greater or lesser extent) these attitudes amongst children. This is consistent to some extent with statistics which show that female participation in sport is increasing, albeit in relatively few sports.

Study Three was designed to test the possibility that children do not hold negative stereotypes of females in sport by presenting the same questionnaire as used in Study Two to a group of adults. While attitudes are subject to change, by and large, the

attitudes we hold as young adults tend to remain in later life (Wachs and Gruen, 1982). It is likely, therefore, that if the questionnaires utilised in Studies One and Two were effective tools for measuring stereotypes, these same questionnaires would elicit stereotypes in a group of adults. It was hypothesised that the results of presenting the questionnaire used in Study Two to a group of adults would provide valuable information to help explain the results obtained so far. For instance, if a group of adults were to demonstrate evidence of negative attitudes then this would suggest that the questionnaire methodology utilised in these studies was effective, and lend support to the possibility that the children do not hold negative stereotypes of females in sport. On the other hand, should the parents not show the expected gender stereotyping, this would suggest that there was something inherently problematic with the methodology used.

Summary

As with Study One the results from Study Two did not show any evidence of widespread negative attitudes or stereotypes towards females in sport. However, the children were presented with descriptions of characters which were solely about their sporting involvement, so unlike Study One, the lack of negative evaluation could not be attributed to the sporting information about the character being overshadowed by other more salient information. Also, the children were able to use a wide variety of rating scales to evaluate the characters so the lack of negative stereotyping could not be attributed to a limited range of possible evaluations.

There appear to be two main possible explanations for the results of Study Two. Firstly, the questionnaire may be an effective tool for measuring stereotypes and attitudes but children may not hold negative attitudes towards females in sport, and this is why no stereotyping was apparent. Secondly, children do hold negative attitudes towards females in sport but the lack of evidence is due to features of the methodology employed. Social judgeability theory (Yzerbyt et al., 1984) might predict that the main problem with the methodology is that children are aware of their own negative attitudes and withhold negative judgements when completing the task required of them. These possibilities are tested further in Studies Three, Four and Five which are described in the following chapters.

CHAPTER FIVE: THE GOLDBERG PARADIGM (MK III): WHAT DO PARENTS THINK OF REBECCA AND DANIEL PLAYING RUGBY AND GOING SWIMMING?

Introduction to Study Three

And the first step, as you know, is always what matters most, particularly when we are dealing with those who are young and tender. That is the time when they are taking shape and when any impression we choose to make leaves a permanent mark.

Plato The Republic

The aim of Study Three was to investigate the stereotypes that adults have of males and females in sport, and to compare the adults' attitudes to those of the children who participated in Study Two. This was to test the possibility that the children who participated in Studies One and Two did not hold stereotypes of males and females in sport.

While the methodology employed in these studies was successful in eliciting some general stereotypes (e.g. that rugby players are more competitive and aggressive than swimmers), no evidence was found to suggest that the children held negative stereotypes of females in sport. It is not possible to tell whether this was due to the children not holding these stereotypes, or whether they did hold such views and there was a problem with the methodology employed.

It may be the case that children today do not hold strong stereotypes of males and females in sport, even though this would be contradictory to previous research and anecdotal evidence. However, it would be expected that adults would still hold strong attitudes towards males and females in sport. This is because, once learned, stereotypes tend to be enduring. The attitudes that people form when young tend to remain in adult life (Wachs and Gruen, 1982). So, if the questionnaire employed in Studies One and Two is an effective tool for assessing stereotypes then it should be able to elicit stereotypes from a group of adults who, it is assumed, still hold strong

attitudes regarding females taking part in sport. On the other hand, if the adults show no evidence of gender stereotyping, this would suggest that there are problems inherent with the questionnaire.

The adults who took part in Study Three were the parents (or guardians) of the children who took part in Study Two. It was assumed, therefore, that the participants in Study Three and Study Two were similar in terms of social economic status and other characteristics which may influence their attitudes. However, it was assumed that the parents differed from the children in that they grew up in an era where strong negative attitudes towards females in sport were even more prevalent. For instance, during the 1960s women were forbidden from participating in many sporting events, e.g. long distance running. However, since then, regulations have changed, and women are now allowed to participate in events such as the marathon. Nevertheless, the attitudes that children learn tend to remain in adulthood. As Plato observed, children are impressionable and what they learn leaves 'a permanent mark'. So, it is possible that the children who participated in Study Two have grown up in a world where society's attitudes towards women in sport have changed from more oppressive attitudes in bygone eras. However, it would be expected that these children's parents would still demonstrate some evidence of negative gender stereotyping: a legacy from their formative years.

Summary

The aim of Study Three was to test the possibility that the children in Studies One and Two did not hold negative stereotypes of males and females in sport. If stereotyping was apparent amongst the adults, this would support the possibility that children today hold different attitudes to those of previous generations. However, if the parents demonstrated no evidence of stereotyping, when prior research suggests very strongly that they should, this would support the possibility that there was some feature of the methodology which markedly influenced the results.

Method

Design

This study comprised a 2 x 2 x 2 design varying the sex of the character in the story (Rebecca or Daniel), the sport of the character in the story (rugby or swimming) and the sex of the participant (male or female). Envelopes containing an introductory letter and two copies of one of the questionnaires were randomly distributed to 120 schoolchildren to take home to their parents or guardians.

Questionnaire Design

The questionnaire used in this study was almost identical to the one used in Study Two. Most of the items, which had been designed for children to understand, were also appropriate for the adults to rate the characters in the vignette. One exception to this was the item in section four which asked the children 'How much would you like him/her as a friend?'. This was considered to be a strange question to be asking adults and so for Study Three this item was changed to 'If you were still at school would you like him/her as a friend?'.

The only other alteration to the questionnaire was the addition of a question in Section Six which asked details about the participants, such as their sex and age. The question 'How many parents/guardians (including yourself) are there in your household?' was included. This was to establish whether a single returned questionnaire represented a 100% return rate from a single-parent household or a 50% return rate from a two-parent household. This was felt to be important as only one questionnaire from each household could be included in the statistical analysis (however, see Participants below). This question was prefixed with the word 'optional' so that any adult who felt that this was too personal a question could omit to answer it.

Instructions to Participants

The envelopes containing the questionnaires also contained a letter to the parents or guardians requesting their help in a research project looking at how people make first impressions of other people. The letter explained that the researcher had been working in their children's school and now wished to compare the children's views to the views of their parents. The letter asked that the adults complete the questionnaires individually and to return them in the envelope provided. A slip was also provided for the parents to request feedback on the results of the study. To assure anonymity of responses the parents were asked to return the slip separately from the questionnaires. Apart from this letter, the instructions on the questionnaire were identical to those given to the children in Study Two.

Participants

Completed questionnaires were returned by 107 parents or guardians (mean age 40.9 years, $SD=5.99$) of children in Year 9 of a comprehensive school in the North East of England. This represents a return rate of 45% of those questionnaires that were distributed.

In total, 70 households returned questionnaires. Of these households, 37 returned two completed questionnaires and 33 returned just one completed questionnaire. In order to fulfil the requirements for an ANOVA, it was originally intended to include in the final analysis all the questionnaires that had been returned singly, and only one of each pairs of questionnaires that had been returned together. This was because one of the assumptions of the ANOVA test is that all observations are independent (Howell, 1992). Including the questionnaires from adults in the same household could be seen as a violation of this assumption. However, after excluding half of the 74 questionnaires which were returned in pairs, it was found that the cell sizes were excessively unbalanced. While the ANOVA test is usually able to account for unequal cell sizes, it was felt that some of the cell sizes were too small to enable a meaningful

analysis. It was, therefore, decided to include all the questionnaires that had been returned. The violation of the independent observations assumption was considered to be less serious than the problem of radically different cell sizes.

Procedure

Envelopes containing two copies of the same questionnaire and a letter of explanation were distributed by the school to 120 Year 9 children who were asked to take them home to their parents or guardians. Two copies of the same questionnaire were sent to each household so that the nature of the study was not immediately obvious to the adults. Each envelope also contained an envelope for the return of the questionnaires. Questionnaires were returned to the school, again via the children, and collected by a teacher who passed them onto the researcher.

Results

The results of the Study Three are similar to those of the first two studies. The adults in this study and the younger sample from Study Two described the characters in broadly similar ways. There is evidence to suggest that the adults completed the questionnaire attentively and that the manipulations were effective. However, as with the children, there was no consistent evidence that the female rugby player was denigrated or that the characters were described in line with negative gender stereotyping.

It should be noted that it is not possible to compare directly the two age groups in the statistical analysis, because the questionnaires were completed by the two samples under very different conditions. In particular, the adults received only written instructions, and completed the questionnaires either individually (or in pairs) rather than in groups. Also, it was not possible to ensure that the adults did not discuss their

answers when completing the questionnaire, as it was with the children. Nevertheless, it is possible to highlight any qualitative differences between the descriptions made by the two samples.

Table 5.1 shows the mean scores for the trait items for each of the four characters. Shaded boxes indicate results that are discussed in the text. It was expected that the female characters would be evaluated more negatively than the male characters, or that the female rugby player would be evaluated more negatively than the female swimmer. However, the mean scores show that the adults tended to describe the four characters in similar ways. For example, the adults tended to see 'strange' as a poor description for all four characters regardless of sex or sport. Similarly, 'athletic' was seen as a good description for all the characters: all the mean ratings for 'athletic' were above 4.00 (*a good description*).

	Male character		Female character	
	swimming	rugby	swimming	rugby
independent	3.67 (1.05) n=24	3.41 (0.84) n=27	3.73 (1.04) n=33	3.52 (0.93) n=21
confident	3.71 (1.04) n=24	3.81 (0.74) n=27	3.73 (1.01) n=33	3.90 (0.77) n=21
competitive	3.83 (0.83) n=23	4.12 (0.65) n=26	4.12 (0.82) n=33	4.62 (0.59) n=21
strange	1.26 (0.54) n=23	1.70 (1.03) n=27	1.36 (0.74) n=33	1.33 (0.66) n=21
attractive	2.41 (1.01) n=22	2.62 (1.06) n=26	2.58 (0.76) n=31	2.71 (0.78) n=21
masculine	2.73 (1.12) n=22	3.08 (1.16) n=26	1.76 (1.15) n=29	2.05 (1.02) n=21

Cell contents: Mean, standard deviation, number of participants
 Rating scale: 1=very poor description to 5=very good description

Table 5.1 Mean scores for Trait Items by Sex of Character and Sport of Character

	Male character		Female character	
	swimming	rugby	swimming	rugby
gentle	2.68 (0.89) n=22	2.28 (0.94) n=25	2.68 (0.98) n=31	2.48 (0.98) n=21
honest	3.86 (0.85) n=21	3.85 (1.01) n=26	3.74 (1.15) n=31	3.38 (1.07) n=21
friendly	3.67 (0.87) n=24	3.77 (0.71) n=26	3.75 (1.05) n=32	3.95 (0.59) n=21
selfish	2.00 (1.11) n=22	2.04 (0.96) n=26	1.58 (0.92) n=31	1.48 (0.81) n=21
feminine	1.74 (1.10) n=23	1.63 (0.93) n=27	2.97 (1.08) n=31	2.57 (0.87) n=21
bigheaded	1.45 (0.91) n=22	2.08 (1.08) n=25	1.31 (0.69) n=32	1.33 (0.73) n=21
hard working	4.36 (0.58) n=22	4.17 (0.92) n=24	4.45 (0.56) n=33	4.29 (0.56) n=21
obsessed	1.96 (0.98) n=23	2.23 (1.03) n=26	2.34 (1.10) n=32	1.81 (0.87) n=21
aggressive	2.13 (1.01) n=23	2.54 (0.95) n=26	2.39 (1.39) n=33	2.05 (0.97) n=21
daring	2.61 (1.08) n=23	3.04 (0.79) n=25	3.15 (1.12) n=33	3.19 (1.12) n=21
unusual	2.09 (1.08) n=23	2.20 (1.12) n=25	1.67 (0.84) n=30	2.71 (1.01) n=21
athletic	4.21 (1.10) n=24	4.23 (0.95) n=26	4.34 (0.83) n=32	4.45 (0.51) n=20
trendy	2.48 (0.90) n=23	2.24 (0.93) n=25	2.35 (0.91) n=31	2.48 (0.87) n=21

Cell contents: Mean, standard deviation, number of participants

Rating scale: 1=very poor description to 5=very good description

Table 5.1 (Contd.) Mean scores for Trait Items by Sex of Character and Sport of Character

Table 5.2 shows the mean scores for the attribution items from Section 2 of the questionnaire (Why do you think Daniel/Rebecca was selected for the summer school?) for the four different vignettes.

	Male character		Female character	
	swimming	rugby	swimming	rugby
She/he was the best rugby player/swimmer in her/his school	4.48 (0.96) n=25	4.41 (0.69) n=27	4.35 (0.85) n=34	4.19 (0.87) n=21
Her/his parents wanted her/him to go	2.79 (1.38) n=24	2.96 (1.26) n=27	2.50 (1.31) n=34	2.86 (1.11) n=21
Her/his parents are good friends with the summer school coaches	2.42 (1.25) n=24	2.23 (1.07) n=26	1.56 (0.93) n=34	1.71 (0.85) n=21
There was a place for anybody who wanted to go	2.00 (1.41) n=23	2.31 (1.09) n=26	2.18 (1.40) n=34	2.00 (1.18) n=21
No other girl/boy in her/his school wanted to go	1.71 (1.04) n=24	1.62 (0.80) n=26	1.56 (0.75) n=34	2.00 (0.95) n=21
She/he was lucky	2.54 (1.32) n=24	2.63 (1.33) n=27	2.15 (0.89) n=34	2.48 (1.21) n=21
She/he is a natural all-round sports player	3.58 (0.93) n=24	3.85 (0.73) n=26	3.85 (0.91) n=33	4.00 (0.63) n=21
She/he tries hard at all sports	3.79 (1.06) n=24	4.19 (0.80) n=26	4.27 (0.84) n=33	4.24 (0.70) n=21
She/he was in the right place at the right time	2.50 (1.18) n=24	3.04 (1.02) n=27	1.94 (1.01) n=34	2.33 (0.86) n=21

Cell contents: Mean, standard deviation, number of participants

Rating scale: 1=very unlikely explanation, 5=very likely explanation

Table 5.2 Mean Scores for Attribution Items (Why do you think Rebecca/Daniel was selected for the summer school) by Sex of Character and Sport of Character.

These results provide no consistent evidence that the female characters were denigrated in comparison to the male characters, or that the female rugby player was denigrated in comparison to the female swimmer. For instance, 'She/he was lucky' was designed as a negative explanation of why the character was selected for the summer school. However, the mean ratings for the four characters on this item were similar, ranging from 2.15 to 2.63, demonstrating that parents did not see any of the characters as luckier than any of the others. This pattern is similar for the positive attribution items such as 'She/he is a natural all-round sports player'. Again, all the characters were evaluated in similarly positive ways: the means were not significantly different, ranging from 3.58 to 4.00.

Table 5.3 shows the mean scores for the four different vignettes for the attribution items from Section 3 of the questionnaire (Why do you think Daniel/Rebecca was selected for the summer School?). As with the previous two sections, these results also demonstrate that all four characters were evaluated in very similar ways. In particular, the characters were all evaluated positively. For instance, on the item 'She/he wants to become a better swimmer/rugby player', all the mean ratings were above 4.50 demonstrating that most of the participants saw this item as a 'likely' or 'very likely' explanation for all the characters. In line with these positive evaluations, the mean ratings for the item 'She/he wanted a holiday in the country' were all below 2.00, showing that the adults saw this item as a 'unlikely' or 'very unlikely' explanation for all the characters.

Table 5.4 shows the mean scores for the four different vignettes for the evaluative items from Section 4 of the questionnaire. Apart from a few exceptions, the characters were evaluated in very similar and generally positive ways. For example, mean ratings for the items 'How much do you like Rebecca/Daniel?' and 'How hard do you think she/he works at her/his homework?' suggest that the characters were all rated as similarly likeable and hard-working.

	Male character		Female character	
	swimming	rugby	swimming	rugby
She/he enjoys physical exercise	4.17 (0.82) n=24	4.15 (0.67) n=26	3.91 (0.79) n=34	4.57 (0.60) n=21
She/he wants to become a better swimmer/rugby player	4.58 (0.72) n=24	4.59 (0.64) n=27	4.53 (0.61) n=34	4.71 (0.46) n=21
It's one step towards her/his ambition of playing/swimming for the county team	4.72 (0.61) n=25	4.56 (0.58) n=27	4.55 (0.62) n=33	4.76 (0.44) n=21
Her/his parents encouraged her/him to go	3.38 (1.01) n=24	3.70 (0.61) n=27	3.12 (0.99) n=33	3.29 (1.15) n=21
Her/his older brother has been to a summer school before	2.54 (1.22) n=24	2.73 (1.00) n=26	2.52 (1.09) n=33	2.57 (1.21) n=21
She/he wanted a holiday in the country	1.92 (1.10) n=24	1.59 (1.01) n=27	1.65 (1.01) n=34	1.57 (.87) n=21
Her/his father is a rugby player/swimmer and she/he wants to follow in his footsteps	2.21 (1.10) n=24	2.58 (0.95) n=26	2.21 (0.93) n=33	2.50 (0.76) n=20
She/he thinks that being sporty will make her/him popular at school	2.25 (1.11) n=24	2.15 (0.77) n=27	1.97 (0.92) n=33	1.95 (0.97) n=21
She/he has some friends who are also going to the summer school	2.70 (1.11) n=23	2.81 (1.06) n=26	2.70 (1.16) n=33	2.67 (0.97) n=21
She hopes to make new friends who also play rugby/go swimming	3.39 (1.03) n=23	3.27 (1.00) n=26	3.41 (1.16) n=34	3.90 (0.83) n=21

Cell contents: Mean, standard deviation, number of participants
 Rating scale: 1=very unlikely explanation, 5=very likely explanation

Table 5.3 Mean scores for Attribution Items (Why do you think Rebecca/Daniel wanted to go to the summer school?) by Sex of Character and Sport of Character

	Male character		Female character	
	Swimming	Rugby	Swimming	Rugby
How much do you like Rebecca/Daniel?	3.56 (0.92) n=25	3.56 (0.80) n=27	3.79 (0.59) n=34	3.57 (0.75) n=21
If you were still at school would you like her/him as a friend?	3.48 (0.87) n=25	3.44 (0.93) n=27	3.68 (0.81) n=34	3.62 (0.80) n=21
How hard do you think she/he works at her/his schoolwork?	3.68 (0.69) n=25	3.78 (0.75) n=27	3.85 (0.76) n=33	3.80 (0.77) n=20
How well do you think she/he will do in her/his exams?	3.60 (0.65) n=25	3.81 (0.79) n=27	3.71 (0.58) n=34	3.55 (0.60) n=20
Do you think she would be a good captain for her team?	3.96 (0.98) n=25	3.81 (0.83) n=27	4.24 (0.74) n=34	4.14 (0.65) n=21
Do you think Rebecca/Daniel should do more work at swimming/rugby or less work at swimming/rugby?	3.32 (0.75) n=25	3.56 (0.75) n=27	3.91 (0.79) n=34	3.60 (0.75) n=20
*How many hours a week do you think Rebecca/Daniel should spend training?	7.91 (4.20) n=23	7.09 (5.48) n=23	11.63 (7.14) n=34	6.03 (2.28) n=19
Do you think Rebecca/Daniel should spend more or less time training?	3.32 (0.85) n=25	3.44 (0.64) n=27	3.44 (0.75) n=34	3.55 (0.69) n=20
Do you think Rebecca/Daniel should spend more or less time doing homework?	3.72 (0.68) n=25	3.67 (0.55) n=27	3.68 (0.64) n=34	3.40 (0.60) n=20
Do you think Rebecca/Daniel should spend more or less time going out with friends?	3.04 (0.79) n=25	2.81 (0.79) n=27	3.15 (0.86) n=34	3.15 (0.67) n=20
Do you think Rebecca/Daniel should spend more or less time watching T.V.?	1.88 (0.73) n=25	2.00 (0.83) n=27	2.00 (0.82) n=34	2.15 (0.88) n=20

Cell contents: Mean, standard deviation, number of participants

Rating scale: 1='negative' evaluation, 5='positive' evaluation

*this item was not answered using a 5 point Likert scale. Participants were asked to write the exact number of hours.

Table 5.4 Mean Scores for Evaluative items by Sex of Character and Sport of Character

	Male character		Female character	
	Swimming	Rugby	Swimming	Rugby
How likely do you think it is that Rebecca/Daniel will play for the county team?	4.00 (1.15) n=25	3.67 (0.88) n=27	4.29 (0.72) n=34	4.10 (0.72) n=20
How likely do you think it is that Rebecca/Daniel will get married?	3.76 (1.01) n=25	4.07 (0.92) n=27	4.26 (0.90) n=34	3.95 (0.76) n=20
How likely do you think it is that Rebecca/Daniel will go to college or university?	3.60 (0.76) n=25	4.11 (0.75) n=27	3.94 (1.01) n=34	3.65 (0.67) n=20
How likely do you think it is that Rebecca/Daniel will have children?	3.72 (1.02) n=25	3.96 (0.94) n=27	4.09 (1.00) n=34	3.65 (0.81) n=20
How likely do you think it is that Rebecca/Daniel will become a well known sports star?	2.60 (1.15) n=25	2.59 (1.08) n=27	2.74 (1.02) n=34	2.85 (0.88) n=20
How likely do you think it is that Rebecca/Daniel will give up her/his sport when she/he leaves school?	2.96 (1.17) n=25	2.48 (1.05) n=27	2.79 (1.15) n=34	3.00 (0.79) n=20
How likely do you think it is that Rebecca/Daniel will represent her/his country at swimming/rugby?	2.60 (1.00) n=25	2.93 (1.24) n=27	3.09 (1.31) n=34	2.80 (1.15) n=20
How likely do you think it is that Rebecca/Daniel will become a swimming/rugby coach?	3.16 (1.07) n=25	3.15 (0.99) n=27	3.26 (1.26) n=34	2.85 (0.99) n=20

Cell contents: Mean, standard deviation, number of participants
Rating scale: 1='negative' evaluation, 5='positive' evaluation

Table 5.4(Contd.) Mean Scores for Evaluative items by Sex of Character and Sport of Character

Table 5.5 shows the mean scores for the four different vignettes for the attribution items from Section 5 of the questionnaire (Why do you think Daniel/Rebecca goes swimming/plays rugby?). As with the previous sections, the four characters were rated similarly regardless of sex or sport. For example, the explanation 'The training

improves her/his good looks' was seen as fairly unlikely for all the characters, with means ranging from 2.06 to 1.88. For the more positive explanations of why the character participated in their chosen sport, such as 'She/he enjoys competitions' and 'She/he does it to keep fit' mean ratings were also similarly high for all the characters.

	Male character		Female character	
	Swimming	Rugby	Swimming	Rugby
She/he does it to keep fit	3.52 (0.92) n=25	3.44 (1.12) n=27	3.56 (0.96) n=34	3.35 (0.88) n=20
She/he enjoys competitions	4.08 (0.76) n=25	4.15 (0.77) n=27	4.03 (0.80) n=34	4.30 (0.57) n=20
She/he enjoys physical exercise	4.16 (0.85) n=25	4.30 (0.54) n=27	3.85 (0.94) n=33	4.30 (0.57) n=20
The training improves her/his good looks	1.88 (0.80) n=24	1.89 (0.97) n=27	2.06 (1.03) n=33	1.90 (0.89) n=21
Being sporty means she/he is popular at school	2.48 (1.00) n=25	2.30 (0.99) n=27	2.18 (1.06) n=34	2.38 (1.20) n=21
A lot of her/his friends also play rugby/go swimming so she/he wants to join in too	3.21 (1.22) n=24	3.19 (1.18) n=27	2.97 (1.17) n=34	2.70 (0.98) n=20

Cell contents: Mean, standard deviation, number of participants
 Rating scale: 1=very unlikely explanation, 5=very likely explanation

Table 5.5 Mean Scores for Attribution Items (Why do you think Rebecca/Daniel plays rugby/goes swimming?) by Sex of Character and Sport of Character

Although included in the statistical analysis, sex of participant was not found to have any consistent effects on the descriptions of the characters and so, for simplicity, the results are shown in Tables 5.1, 5.2, 5.3, 5.4 and 5.5 for all participants together.

A comparison of the mean scores in Table 5.1 and Table 4.7 demonstrates that the adults and the children described the four characters in broadly similar ways. One exception to this is the adjective 'trendy'. The children saw the female swimmer and

the male rugby player as the most trendy ($M=2.53$ and $M=2.86$ respectively), whereas the adults saw the male swimmer and the female rugby player as the most trendy ($M=2.48$ and $M=2.48$).

As expected the adults described the male characters as more 'masculine' than the female characters, $F(1,90)=18.42$, $p<0.001$. As in Study Two, no interaction was found between sex of character and sport of character.

Similar results were found for 'feminine'; the female characters were rated by the adults as significantly more feminine than the male characters, $F(1,94)=29.58$, $p<0.001$. However, unlike Study Two, there was no significant interaction between the sex of the character and the sport of the character. While the female swimmer was rated as the most feminine of the four characters, this character was not rated significantly more feminine than the female rugby player.

In Study Two, the female rugby player was rated by the children as significantly more unusual than the other characters (see Figure 4.2). As shown in Figure 5.1, a significant interaction between sex of character and sport of character was also found with the adult sample, $F(1,91)=4.89$, $p<0.05$. A Scheffé post-hoc analysis revealed that the female rugby player was rated as significantly more unusual than the female swimmer. As with Study Two, no main effects or interactions were found for 'strange'. This would support the interpretation that the adults also viewed the female rugby player as unusual in terms of being a rarity rather than as an oddity.

A further result which may be indicative of gender stereotyping is the significant interaction between sex of character and sport of character for the item 'obsessed', $F(1,94)=4.06$, $p<0.05$. As Figure 5.2 shows, the female swimmer was rated as the most obsessed and the female rugby player as least obsessed. However, post-hoc analyses revealed that none of the mean ratings for the characters were significantly different from any other.

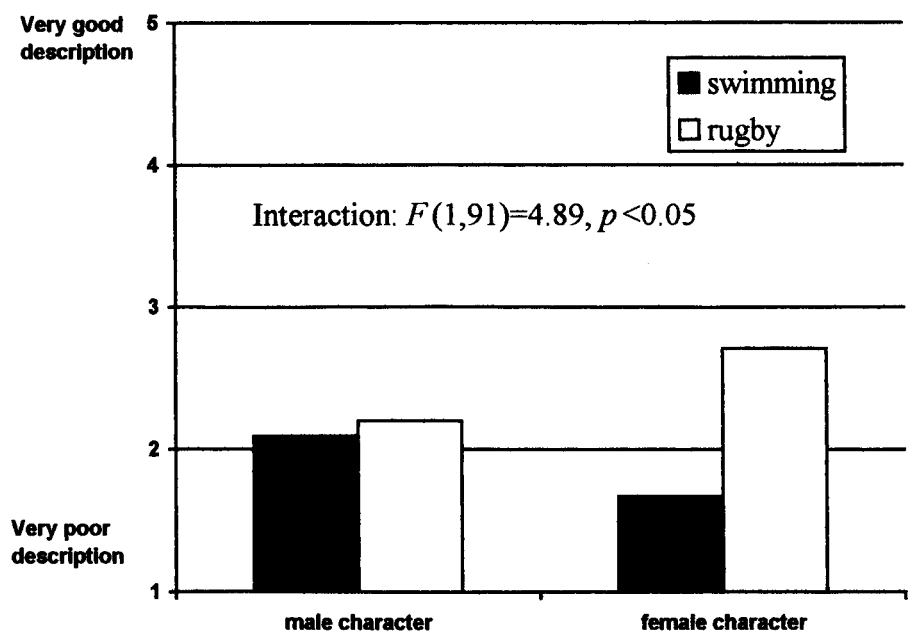


Figure 5.1 Mean ratings for 'Unusual' by Sex of Character and Sport of Character

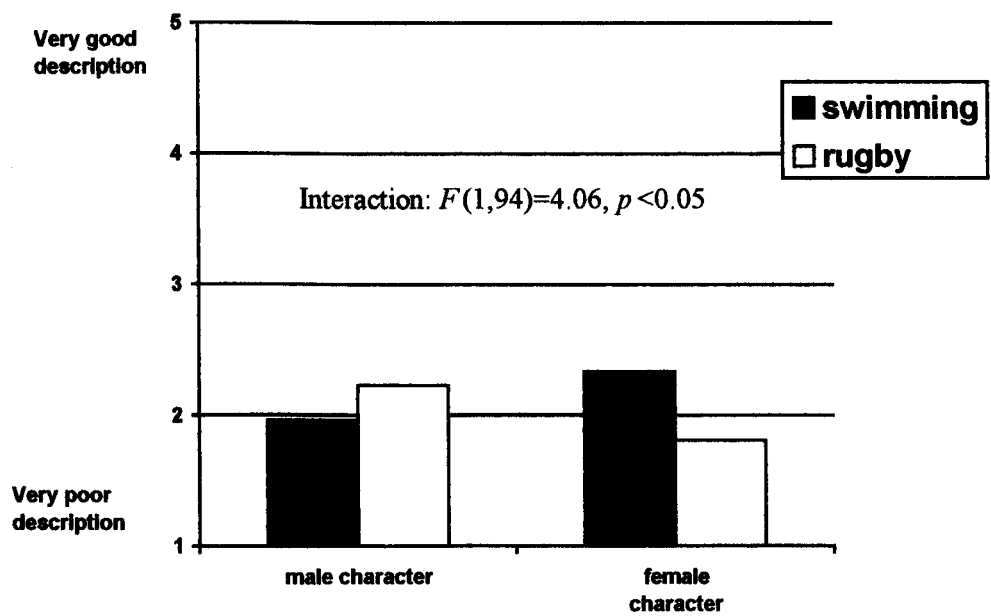


Figure 5.2 Mean ratings for 'Obsessed' by Sex of Character

A further result which might suggest evidence of stereotyping was found for the only item which was not answered using a five point Likert-type scale. This was the item asking the participants how many hours the character should spend training each week (see Figure 5.3). The interaction between sex of character and sport of character was found to be significant, $F(1,91)=4.66, p<0.05$. On average, participants suggested that the female swimmer should do the most training each week, and the female rugby player the least. A Scheffé post-hoc test revealed that the difference between the ratings in these two conditions was significant.

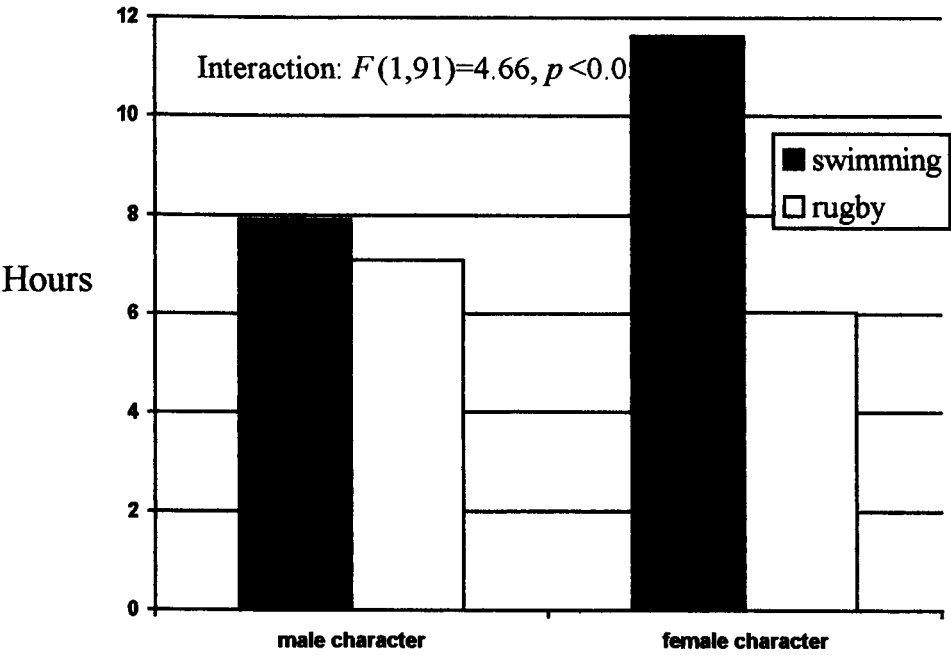


Figure 5.3 Mean estimates for 'How many hours should Rebecca/Daniel spend training each week?' by Sex of Character and Sport of Character

Apart from the results shown in Figures 5.1, 5.2 and 5.3, the results provide no consistent evidence that the female rugby player or the female swimmer were evaluated in line with gender stereotypes. However, as with the previous two studies there was evidence that the various characters were evaluated by the adults in line with gender stereotypes or with stereotypes of swimmers and rugby players. For instance, the rugby players were rated as more competitive than the swimmers,

$F(1,95)=6.76$, $p<0.05$, and interestingly, the female characters, irrespective of sport, were rated as more competitive than the male characters, $F(1,95)=6.89$, $p<0.05$.

The results for 'How likely do you think it is that Rebecca/Daniel will go to university?' may provide some evidence of sport stereotyping. For this item a significant interaction was found between sex of character and sport of character $F(1,98)=5.01$, $p<0.05$. A LSD post-hoc test revealed that the adults rated the male rugby player as significantly more likely to go to university than the male swimmer. This may reflect a stereotype which associates rugby playing with male university students.

A main effect of sex of character was found for the item 'bigheaded', with the male characters rated as significantly more bigheaded than the female characters, $F(1,92)=5.73$, $p<0.05$. Similarly, the male characters were rated as more selfish than the female characters, $F(1,92)=6.12$, $p<0.05$. While these results may reflect stereotypes of males and females, contrary to expectations, there is no evidence to suggest that the female taking part in rugby was evaluated in any different way from the female taking part in swimming.

Results for the attribution items were similar to those found for the trait items. In general, while there were a few main effects and interactions between sex of character and sport of character, there was no evidence to suggest that any of the four characters were being attributed with different reasons for their selection or their participation in sport. In particular, there was no consistent evidence that the female rugby player was attributed with less desirable motives or explanations.

The results for only one of the attributions explaining why the characters wanted to go to the summer school showed a significant interaction between sex of character and sport of character. A Scheffé post-hoc analysis confirmed that 'She enjoys physical exercise' was rated as a more likely explanation for the female rugby player than for the female swimmer, $F(1,97)=6.43$, $p<0.05$. However, other similar

attribution items, such as 'She enjoys competition', which was also designed as a stable, internal and positive attribution (see Table 4.5), did not reflect this finding.

Discussion

In summary, the results from Study Three are very similar to those from Studies One and Two. There is no reason to doubt that the parents understood the questionnaire or that they completed it attentively, particularly as there is evidence that the adults demonstrated sport stereotypes and also some gender stereotypes. However, as with the previous two studies, the results show no consistent evidence that female sporting characters in the vignettes were denigrated or negatively stereotyped.

The results for the items 'feminine', 'masculine', 'unusual', and 'strange' are similar to those from Study Two and suggest that the participants did understand the vignettes and the rating scales and that they completed the questionnaire with some attention. However, none of these results appear to indicate a negative evaluation of the female sports participants or of the female rugby player in particular.

It is interesting that the parents and children differed in their descriptions of the characters on the item 'trendy'. The children saw the female swimmer and male rugby players as the most trendy. Perhaps for the children those people who take part in 'appropriate' activities are 'cool' or 'fashionable'. However, the parents may be more aware of conventions and they describe the female rugby player and male swimmer as 'trendy' because they are breaking traditions and are more 'modern'. This highlights a particular problem of this methodology; different people may interpret the items in different ways. 'Trendy' could be interpreted as 'modern' or as 'fashionable'. This could explain why the parents and children differed in their views of the characters when deciding how good a descriptor the word 'trendy' was.

The results for 'obsessed' and 'How many hours a week do you think Rebecca/Daniel should spend training?' may be indicative of a negative evaluation of the female rugby player. This may seem counter-intuitive as the female rugby player was rated as the least 'obsessed' and this would normally be interpreted as a negative trait. However, it may be that, in the context of competitive sport, obsession is a necessary and desirable trait. This may be an even more desirable trait in a sport such as swimming where young participants often have to train for several hours each day before and after school. This training requirement is reflected in the parents' judgements of how many hours the characters should train each week. While the swimmers were expected to train more than the rugby players, this difference was only significant for the female characters. Therefore, the expectation that the female swimmer should train considerably more each week than the female rugby player cannot be explained purely by a stereotype of swimming as a sport which requires long hours of training. The parents' judgements that the female swimmer should train much longer each week than the female rugby player could be interpreted as a withdrawal of encouragement or approval. That is, for a sport which is perceived as appropriate for females, the parents may view long hours of training and obsession as acceptable, without any negative connotations. In this way, the female swimmer may be given more approval than the female rugby player. On the other hand, this evidence is tentative and is not supported by any other consistent results from the other items. Swimming is a sport which is generally viewed as acceptable for female participants and it also requires long hours of training. These facts alone may account for the results for the items 'obsessed' and 'hours of training', without resorting to negative gender stereotyping as an explanation.

Apart from a few exceptions, the characters were described in similar ways. However, the male characters were described as more selfish and bigheaded than the female characters. Also, the male rugby player was rated as the more likely to go to university than the other characters. These findings seem to reflect general sport stereotypes or gender stereotypes. There was no consistent evidence to suggest that the female rugby player or the female swimmer were denigrated or negatively stereotyped.

Overall, apart from a few exceptions, the results from Study Three are very similar to those from Studies One and Two. In all three studies, evidence was found to suggest that the participants did evaluate the characters in line with sport stereotypes or gender stereotypes. However, there was no consistent evidence to suggest negative gender stereotyping of the female sporting characters. The questionnaire did elicit some stereotyping but not in the expected ways. In particular, the results of Study Three were strongly expected to demonstrate that the parents held strong attitudes towards males and females taking part in 'appropriate' sports. This expectation was not fulfilled.

The results of Study Three call into question the effectiveness of the methodology employed in this research because there were strong expectations that the questionnaire would be effective in eliciting from parents negative stereotypes of females in sport. One of the problems with this study is that there was no control over the circumstances in which the parents completed the questionnaire. In the classroom it was possible to prevent the children from becoming aware of the other conditions. Although only copies of the same questionnaire were sent home to each household, it is possible that parents of different children discussed the research and realised that they had different questionnaires.

A further problem with this questionnaire is the assumption that the participants do not know the exact nature of the study. They know that they are being asked to evaluate a character but they do not know that the study is investigating people's attitudes towards males and females in sport. While this may be true for the conditions where the character is described as participating in sports which are sex-appropriate (i.e. male swimmer, female swimmer and male rugby player) this assumption may be violated in the condition where the female character is described as playing rugby. This description may be quite surprising to the participants and draw their attention to issues surrounding gender and sport. In this way, the parents who were asked to evaluate a rugby-playing female may have had strong suspicions regarding the purpose of the study. This may have led to these participants, in

contrast to those in the other conditions, responding in socially desirable ways. This might explain why the female rugby player was not evaluated negatively.

Two possible explanations for the results of Studies One and Two were discussed in Chapter Four. The first main explanation was that children do hold strong stereotypes of males and females in sport and that there was some feature of the methodology that was employed which resulted in the apparently egalitarian attitudes which were reported. The second main possibility was that the questionnaire methodology which was employed is an appropriate tool for assessing stereotypes, and that no stereotypes of males and females in sport were found because children do not hold such stereotypes. The results of Study Three lend support to the former explanation.

It is not possible to discount the possibility that stereotypes of males and females in sport have changed. However, the results of Study Three suggest that there may be some problems with the Goldberg paradigm as a method for assessing stereotypes. The rationale behind using a disguised attitude measure was to avoid the participants responding in socially desirable way. However, according to social judgeability theory (Yzerbyt et al., 1994), it is possible for participants to withhold their 'true' attitudes even if they are not aware of the exact nature of the study. That is, before making a judgement about another person, the participant has to feel confident that they are in a position to judge. Confidence will be influenced by various factors or 'levels of adequacy'. One of these levels of adequacy relates to cultural norms. That is, we are all aware of cultural norms and respect these norms when making judgements of other people. Yzerbyt et al. suggest that one of the norms prevalent in Western culture says that it is wrong to make judgements about other people if you do not have any more information about a person other than their category membership.

The participants in Studies One, Two and Three may have very strong attitudes towards males and females in sport but, when presented with the vignette and a list of questions asking them to make judgements of the characters, felt that it would be wrong to express these attitudes. So, while the exact nature of the studies may not have been apparent, the participants may have realised that they were being asked to

make judgements of characters about whom they knew virtually nothing. Informing participants that the study was looking at 'how people make first impressions of other people' may have exacerbated this problem. According to social judgeability theory this would result in participants withholding their 'true' judgements.

If the social judgeability explanation is correct, then it would appear that if the participants were respecting norms then these norms must be 'stereotype-specific'. That is, in all three studies, participants expressed some stereotypes. In particular, the participants had strong stereotypes of the various sports which the characters were involved in. For instance, the participants in Study Two viewed the rugby players as more competitive and aggressive than the swimmers. It may be that there are no norms which suggest that it is wrong to stereotype sports, or that there are such norms but participants did not respect them in the context of this study. It is possible to imagine a different context where participants might withhold their judgements about different sports. What is clear from Studies One, Two and Three, is that the participants did not denigrate or negatively stereotype the female rugby player or female swimmer. From these studies it is not possible to tell if this is due to the participants not holding stereotypes of males and females in sport or whether they did hold such stereotypes but, in respecting norms, they withheld their judgements.

Apart from the social judgeability explanation of these results, there are other features of the methodology which may have resulted in the apparent lack of stereotyping. For instance, the traits used may not have meant the same to the participants as they did to the researcher. As previously mentioned, items such as 'trendy' may have different connotations and be interpreted in different ways. Even if participants do describe a character as obsessed and competitive, it is not possible to conclude that this is necessarily a negative evaluation. A further problem with the items is that the participants could only endorse, to a greater or lesser extent, the items that were presented to them. It is possible that the participants did have very strong attitudes but were unable to express them using the items which were presented to them. Dimensions which appear salient to the researcher are not necessarily salient to the participants.

The results from the first three studies are consistent with the possibility that children do not hold strong stereotypes of males and females in sport. However, the results are also consistent with a social judgeability theory interpretation which suggests that even though participants are not aware of the exact nature of the study they are aware of social norms which specify that it is wrong to make judgements about people unless enough is known about them.

Previous research and anecdotal evidence from teachers would suggest that it is highly unlikely that children do not hold strong attitudes towards males and females in sport. It is clear, though, that the questionnaire methodology utilised in Studies One, Two and Three was not effective in eliciting these attitudes. In order to explore these issues further, it was decided to utilise a different methodology - a semi-structured interview - in Study Four.

As well as aiming to illuminate the findings from the previous three studies by employing a radically different methodology, the interviews were also used to inform the design of the questionnaire employed in Study Five. As mentioned previously, a problem with the questionnaires used in the first three studies was that participants could only endorse items which had been presented by the researcher. It was hoped that the interviews could be used to generate items which were meaningful to the children, by asking them to describe their attitudes and feelings in their own words.

Summary

As with the previous two studies, the results from Study Three did not show any widespread negative stereotypes of females in sport. The results from the first three studies are consistent with two possible explanations. Firstly, contrary to the findings of previous research, children may not hold negative stereotypes of females in sport. Secondly, children do hold negative attitudes and stereotypes but, in line with social judgeability theory, they withhold their 'true' judgements because they do not feel confident enough to do so. Whether people feel confident or not to make judgements

depends, in part, on social norms regarding the appropriateness of evaluating other people on the basis of little information.

Study Five, described in Chapter Seven, utilising a methodology similar to that employed in Studies One, Two and Three, attempted to manipulate children's feelings of confidence in order to test the predictions made by social judgeability theory. Study Four, described in the next chapter, utilised semi-structured interviews to investigate further children's attitudes towards males and females in sport, and aimed to establish whether children do hold negative stereotypes of females in sport.

CHAPTER SIX: THE INTERVIEW STUDY: A DIRECT ASSESSMENT OF CHILDREN'S ATTITUDES AND STEREOTYPES

Introduction To Study Four

As in Studies One and Two, the overall aim of Study Four was to investigate children's attitudes towards males and females in sport. However, following the results from the previous three studies, a different methodology - semi-structured interviewing - was employed.

There are various possible reasons why the children in Studies One and Two did not show any negative attitudes towards females in sport. The results of Study Three suggest very strongly that, rather than the children not holding strong attitudes, there may be some features of the methodology which could explain the results from Studies One and Two. Study Four was designed to try to overcome some of the problems created by the features of the 'Goldberg Paradigm'.

Study Four, while also investigating children's attitudes towards males and females in sport, was different from Studies One and Two in several crucial ways. For instance, the children were not constrained in the types of responses they could give. Studies One and Two only permitted children to endorse to a lesser or greater extent the items which had been generated by the researcher. While the items in Study Two were more numerous and more varied than in Study One, it is still possible that the children did hold strong negative attitudes towards females in sport, but that they were unable to express these views with the items that were presented in the questionnaire. In Study Four, the children were free to respond in their own words.

As well as learning more about children's attitudes towards the sexes in sport, it was hoped that the data collected from Study Four could be used to inform the design of further questionnaires. For any questionnaire to be effective, it is vital for the participants to understand the attitudes it aims to investigate and the language used to describe these attitudes. By allowing the children in the present study to express their

attitudes in their own words it was possible to learn more about the specific language that they use to describe their feelings and beliefs.

In Studies One and Two each child read about only one character. However, during the interviews, the participants were asked to consider general questions (e.g. Are there any sports that you think girls shouldn't do?) as well as more specific questions (e.g. How would you feel if you were told that you were going to play netball in your next P.E. class?). It may be that the situation presented to the children in the questionnaires was too abstract and that they found it difficult to express their views on a hypothetical situation. It was hoped that this difficulty would be overcome by presenting scenarios to which the children could more easily relate. On the other hand, the children may have found that the scenario described in the vignette was easy to relate to and that it was easy to form opinions about the character, but they may have felt reluctant to express views about an actual named person. This is discussed below with respect to social judgeability theory, but it can be noted here, that in the interviews, it was hoped to remove this inhibition by asking more general questions which did not relate to any named (fictitious or otherwise) individual.

Care was taken in Studies One and Two to reassure the children that there were no right or wrong answers to the questions asked. Nevertheless, the children may have experienced some degree of evaluation apprehension, especially as they completed the questionnaires under conditions similar to those of an examination. If a child did suffer evaluation apprehension it is likely that this would have influenced their responses. It was hoped that the interviews would be perceived by the children as being very different from an exam and that apprehension would be reduced. During the interviews it was possible to ask questions such as 'What do you think your classmates would say about a boy playing netball?'. In this way it is still possible to learn about children's attitudes towards the sexes in sport but pressure is removed from an individual child who may be self-conscious about expressing their own views.

An important feature of semi-structured interviews is that it is possible to follow up issues and to ask the participants to embellish particular responses. In this way, a participant can be encouraged to express complex attitudes and to qualify their responses in ways which are not possible using a Likert-type scale.

In short, the main aim of Study Four was to investigate children's attitudes towards males and females in sport. Of particular interest was children's perceptions of the appropriateness of various sports for males and females. It was hoped that a semi-structured interview would be more successful at investigating these attitudes than the questionnaires used in Studies One and Two.

While the primary aim of Study Four was to learn more about children's attitudes and the stereotypes they hold, a further aim was to illuminate the results of the previous three studies. As mentioned in Chapter 5, social judgeability theory (Leyens et al., 1994) may explain why, in contrast to previous research, the children in Studies One and Two showed no evidence of negative gender stereotyping. Leyens et al. suggest that before making judgements of other people we have to feel that we are in a position to judge. Whether we feel in a position to judge or not will depend on factors such as the amount of information we have about the person, or the type of information we have about the person. These authors suggest that, as social beings, we respect social norms when making judgements about people. One norm which may influence our judgements of other people is one which says that it is wrong to judge other people when the only information we have about them is categorial, e.g. whether they are male or female. Thus, when people feel that they only hold categorial information about a target, they will often avoid making a judgement and instead provide, in Leyens et al.' terminology, an 'omnibus response'. However, if individuating information about the person is provided, people feel confident that they are able to make a judgement and will use this information to make a judgement. Leyens et al. propose that stereotyping will occur when people feel confident that they are in a position to judge a target individual but when in fact they only have categorial information. This was the basis for the studies described in Chapters Three, Four and Five that utilised the Goldberg paradigm. Participants were presented with a story

about a target individual which was specifically designed to give the impression that they knew a lot about the person, when in fact they knew very little. However, the participants showed no evidence of stereotyping of the sexes in sport. The possibility that children do not hold strong views of males and females in sport has been discussed previously. It seems more likely that it is some feature of the questionnaire or its administration that is making the children reluctant to make stereotypic judgements of the target individuals. Perhaps they are respecting the social norm that says it is wrong to make judgements of people when you have little information about them.

To explore these possibilities it was decided to carry out semi-structured interviews with similarly aged children as those that had completed the questionnaires previously. This change from an indirect method to a direct method of assessing stereotypes appears to be contradictory to the rationale provided for using the 'Goldberg paradigm' in the first three studies. However, if the predictions made by social judgeability theory are correct, there are features of the interview situation which may make children more likely to stereotype the sexes even though they are aware of the attitudes being assessed. Firstly, it was hoped that in an interview situation it would be easier to convince a child that there are no right or wrong answers and that they are not being tested. This was emphasised during the administration of the questionnaires but the children may still have perceived that they were being assessed in some way. The fact that many of the children expected to put their name on the questionnaire supports this suggestion.

If they experienced 'evaluation apprehension' they may have been more likely to answer in the way they felt they should (i.e. follow social norms). The second feature of the interview situation that may make children more confident about making (stereotypic) judgements is that it is possible to talk in general terms about males and females in sports. It is not necessary to consider specific target individuals as was the case in the questionnaire studies. It may be socially acceptable to make generalisations of groups of people but unacceptable to make judgements of one person based on their group membership. This point may be particularly important in

stereotyping research today. It remains to be seen whether or not children hold strong views of the sexes, but given the apparently anomalous results from the present research, the more important issue may be *when* do children express stereotypic views and *when* do they withhold them.

To sum up, the interviews were carried out with two aims. The primary aim of Study Four was to investigate children's attitudes towards males and females in sport. The second aim was to illuminate the results of Studies One, Two and Three, and to explore the predictions made by social judgeability theory.

Method

Design

A semi-structured interview schedule was employed to investigate the attitudes of 12 and 13 year-old boys and girls towards males and females in sport.

Participants

The participants were 12 and 13 year-olds from a local comprehensive school in a large city in the North East of England. In total, 13 boys and 12 girls were interviewed.

Pre-pilot Study and Pilot study

A pre-pilot study was carried out in a different school from that used to recruit the participants for the main study. Two 11-year old children, one male and one female,

were interviewed about various, general sporting issues to establish which kinds of topics might be appropriate to include in the main interviews. Following these interviews, a draft interview schedule was developed and piloted with a girl who attended the school from which the main sample was drawn. These pilot interviews informed the development of the final interview schedule.

The Interview Schedule

The purpose of this study was similar to that of Studies One and Two in that it aimed to assess children's attitudes towards males and females in sport. However, Study Four intended to fulfil this aim while placing less constraint on the kinds of answers that the children could give than in the previous studies. The interview schedule was designed to be a starting point from which discussion of other issues could develop. However, a clear set of questions was developed, along with appropriate prompts, to aid those children who were reticent to embellish their answers. Care was taken to avoid any questions which required 'yes' or 'no' answers or which were too abstract.

Each interview started with general questions such as 'Can you tell me what sort of things you do in your spare time?' and 'Do you do any sports?'. These questions were intended to be straightforward and 'non-threatening' for the participants, who may have found the situation unusual. The questions then moved onto the issue of gender and whether the participant played sports in same-sex or mixed-sex groups. In order to gauge attitudes concerning the appropriateness of different sports, the children were asked what their reaction would be to playing either netball (male participants) or rugby (female participants) in their next P.E. class. The aim was to discuss with each participant a concrete example to which they could easily relate. What their class-mates' reactions would be was also discussed and whether this would be different from their own reaction. Each participant was also asked what their reaction would be if a classmate decided to take up rugby (female classmate) or netball (male classmate) at a club outside of school. The full interview schedule for a male participant is shown in Figure 5.1. For female participants the schedule was the same

except that they were asked what they felt about girls playing rugby or participating in boxing. Note that the interviewer was free to omit questions, add questions and change the order in which the questions were presented depending on the progression of each individual interview.

What sort of things do you do in your spare time?

- sports?

Who do you do these sports with?

- same-sex? mixed-sex?

I am particularly interested in the sports that young people do. Can you tell me, do you think boys or girls are more interested in sport, or is it about the same?

- why?

What sports do boys typically do?

Are there any sports that you think boys shouldn't do?

- why not?

What would the reaction be of the boys in your class if you went down to your next PE lesson and the teacher said "Okay boys today you are going to do netball"?

- pleased?

- what sort of thing would they say?

- why?

What if one boy decided he really liked netball and joined a club? How would the boys in your class react?

- tease him? call him names? encourage him?

- why?

- what would your reaction be? different? the same?

What if he became really good and started winning trophies and medals?

- would your reaction be the same?

- jealous? proud?

Figure 5.1 : Interview schedule for male participants

(female participants were asked to consider girls participating in 'rugby' or 'boxing')

Procedure

The participants in this study attended the same school as the participants in Study Two. For the interview study, the youngest year group in the school was selected to ensure that none of them had been involved in the previous study. It was, however, not possible to ensure that the participants did not have a sibling or a parent who had been involved in either Study Two or Three.

Two classes of children (approximately 60 pupils) took a letter home explaining the purpose of the study, and which gave parents the opportunity to withdraw their child from the study. No parents objected and the class teacher selected pupils to be interviewed at the start of each lesson. The interviewer asked that approximately equal numbers of males and females be selected from a wide range of academic abilities and sporting involvement. On occasion, participants specifically asked the teacher to take part in an interview. In total, 13 boys and 12 girls were interviewed over a 5 week period.

The majority of the interviews took place in the Departmental staff room close to each child's classroom. However, due to interruptions from staff requiring resources from this room during class time, some interviews took place in an empty classroom. All interviews took place during geography lessons and took between 10 and 20 minutes to complete.

The interviewer explained that he was from the University and was carrying out research looking at the sports that children do and what children feel about doing different sports. It was explained that the interview was going to be recorded because it was not possible for the interviewer to write everything down. However, the tape-recorder was placed close to the participant who was told that they could switch it off at any time if they felt uncomfortable or if they wished to say something but not have it recorded. All participants were assured that everything they said would be held in confidence. Participants were also encouraged to ask the interviewer any questions that they had.

The interviewer, and author of this thesis, was a 26 year-old white male with a Scottish accent. As the interviewer was not local to the North East of England there was occasional confusion over some dialectical terms. However, when confusion arose it was usually possible to encourage the participants to embellish their responses.

It was important not to alienate those children who were not interested in sport but who still had strong feelings or opinions about their non-participation. Had the interviewer appeared to be sporty himself, the children may have been induced to emphasise the positive aspects of sport. So as to not appear particularly sporty (or too much like a teacher), the interviewer took care to dress casually but not in any sportswear such as a tracksuit or training shoes.

All interviews were tape-recorded and later transcribed by the interviewer, with one case (female) being eliminated from analysis due to the poor quality of the recording. The transcripts were supplemented by notes taken during and after each interview. To maintain the anonymity of the participants, when transcribing the interviews, the researcher used a pseudonym for each interviewee.

In general, the participants were interested and keen to take part in the interviews, possibly because it meant time away from class. While the majority of participants answered the questions freely, most were reluctant to 'embroider' their responses and appeared to expect the interviewer to guide the direction of the interview. This may be characteristic of interviewing children who are used to giving 'the right answer' rather than their own opinions (Breakwell, 1995).

Data Analysis

The data analysis employed was similar to that described by Huberman and Miles (1994). Their method involves an iterative process consisting of data display, data reduction and the drawing and verification of conclusions.

To display the data, a word-processing package was used by the interviewer to transcribe the interviews directly onto computer. This same word-processing package was employed in the coding of the data. Relevant segments of text were copied from the interview transcripts and collated into new files with other segments on similar themes. The themes that were used to collate segments of text were those based on the researcher's expectations and predictions (formulated following the results of Studies One, Two and Three) as well as those that emerged during and after the interviews.

The segments of text used for analysis were mostly sentences or phrases uttered by the participants but in some cases the relevant segment also included questions or comments from the interviewer. A particular segment could be copied to more than one file if it was relevant to more than one category. In this way the data were reduced into more easily-manageable files which contained segments of text on related topics. Preliminary analysis of these files was then used to inform further coding and categorisation of the transcripts. This process of data coding and categorisation was repeated until further categorisation produced no further meaningful themes. The final step of analysis was to generate explanations of the data and to draw conclusions.

The themes used to reduce the data in the first instance were derived from the main research questions. That is, in investigating children's attitudes towards the sexes in sport it was hoped to identify any activities which were considered inappropriate for either sex. Of prime importance were the reasons given to explain why these activities might be inappropriate. Furthermore, while the analysis concentrated on what the children expressed, the ways in which they expressed their views were also

studied. Figure 5.2 illustrates the coding scheme that was developed to guide the process of analysis. Once the data had been coded and categorised according to these key themes, analysis continued using themes which emerged from the data.

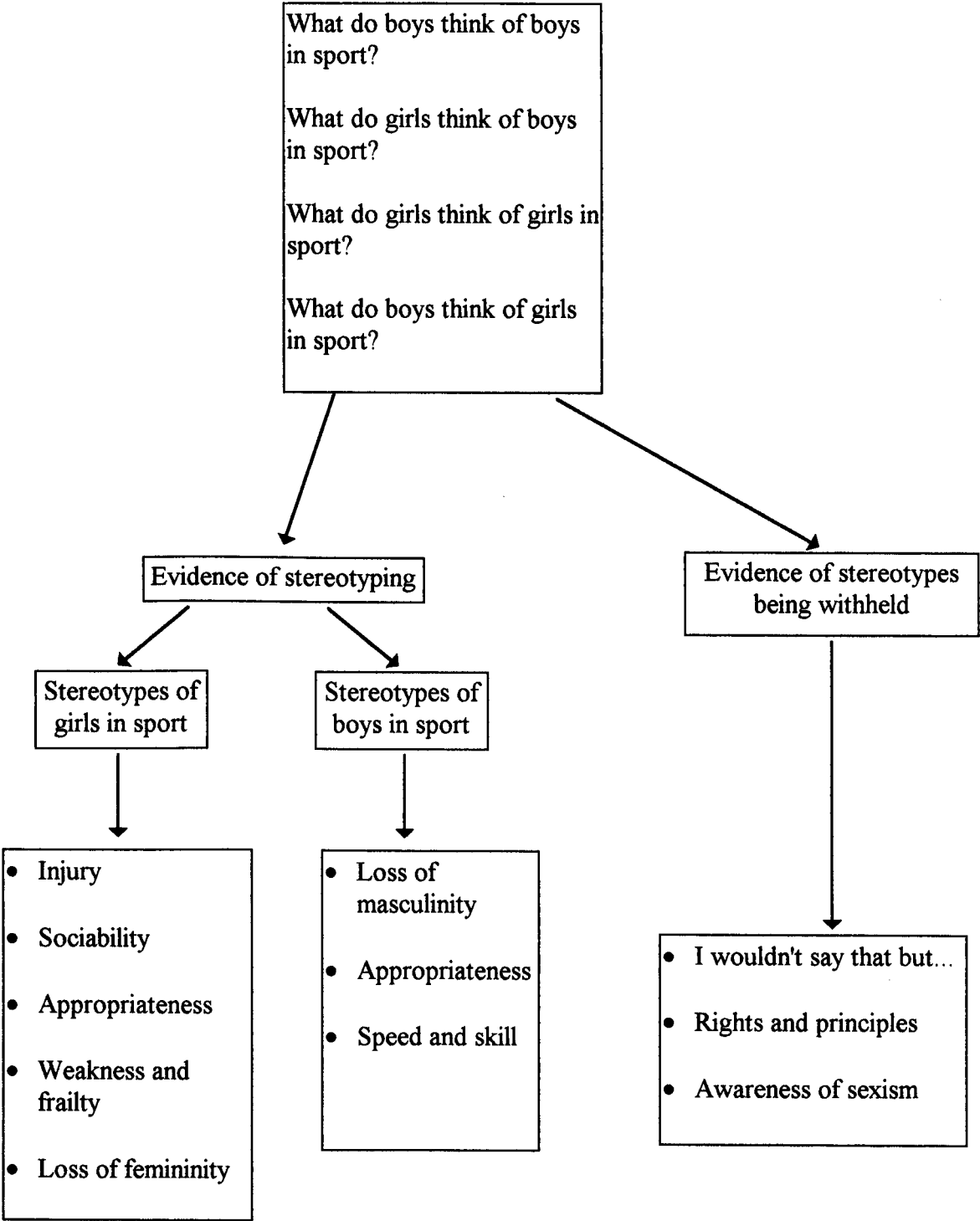


Figure 5.2 Scheme of coding categories used for study four data analysis

Results and Analysis

Overview: Sports Participation

The vast majority of the children reported that they took part in some kind of sport during their spare time. Only one boy did not list any sports amongst his hobbies. Girls reported taking part in more activities than the boys, and a far wider range of activities such as swimming, cycling, dancing, athletics, badminton, tennis and football. The school has a girls' football club and two of the girls interviewed played for one of the teams. Boys, on the other hand, reported participating in far fewer sports, but they were involved in these sports up to a higher level. For instance, several of the boys reported playing football for the school team or for youth teams. Other sports mentioned included cricket, swimming and running. Sport seemed to play a much bigger part in the boys' spare time as there were few other activities mentioned apart from watching TV and playing computer games. This is in contrast to the girls who listed non-sporting activities such as reading, drawing, shopping, going to the cinema and playing musical instruments.

It appears that the boys and girls in this study do conform to traditional views of the sexes with regards to sports participation, with the exception that girls did report playing football to a competitive level. Boys may report fewer non-sporting activities and over-emphasise sports participation because of social desirability effects. Previous research has encountered similar problems with self-report measures of sport participation (e.g. Wright and Macleod, 1992). A problem which arose in the present study was that, on occasion, in asking children about their sport participation, the children had very different views from the interviewer of the intensity or frequency of an activity. For instance, some children reported running as a sport they took part in. Further probing established that what was often meant was 'running around' (e.g. in the street or playground) rather than formally structured athletics.

Attitudes towards females in sport

The results from Study Four demonstrate that, in contrast to Studies One and Two, the children did hold very strong attitudes towards males and females in sport. In particular, the children had very strong ideas about which sporting activities were suitable for males and which were suitable for females. When asked what sports they thought boys and girls usually did, football, rugby and cricket were identified as boys' sports and netball and hockey as girls' sports. Notably, some children pointed out that girls also play football. This can be attributed to the success of the school's girls' football team, which several of the interviewees mentioned.

While nearly every participant believed that people should be able to take part in whatever sport they wanted to (a point discussed below), there was general resistance to the idea of girls taking part in sports such as rugby and boxing. One boy even suggested that girls do not like sport.

James : Cos more boys like sports girls just like to walk around talking

When asked why girls should not do these sports, one of the reasons put forward was that girls could get injured. James suggested that girls should not play rugby because they can get hurt. Similarly, Pete suggested that girls should not box because "it's too rough for them, just in case they get hurt too much". When asked if boys do not get hurt as well, Pete admitted that men also get hurt and get put into hospital, but then suggested that girls could not train well enough for boxing:

Pete : It's they're not like built for it they're not like brought up like men are.....they're not strong they don't train and all that.

However, in summarising his feelings, Pete said that it is up to girls if they want to get hurt, but he still did not think that they should box. It is interesting to note that Pete attributes the differences to males and females, at least in part, to the way they are brought up rather than to innate biological differences.

Elaine reflected Pete's view that girls might be more susceptible to injury. She suggested that if her female classmates were confronted with rugby in a games lesson they might say "Oh god! What happens if I break a leg" or "What happens if I get really hurt". However, while several of the female participants demonstrated resistance to girls playing rugby, their views were considerably less strong than those expressed by the male participants.

The issue of injury was also important for James who suggested that girls should not play American football because it was "dangerous" and "people ...get jumped on by dead heavy men". He attributed the growing popularity of women's football to the development of equipment which was designed to reduce injury or pain.

James: Cos in them days they had dead hard balls and they couldn't header it without hurting yourself.....[Interviewer: Is it different now?] Well the balls aren't as hard...and people wear shin-pads now, so you can't get hurt.

James's responses revealed another problem with researching this topic. When asked about girls playing rugby or football, several of the children assumed, at first, that the interviewer was asking about girls playing in boys' teams. It was, therefore, important that the researcher distinguished between responses that reflected disapproval of girls playing rugby at all and the (possibly) less extreme responses which only referred to a disapproval of girls playing rugby with boys.

Apart from the specific danger of injury, there were other more general reasons given to explain why girls should not take part in particular sports. Either these sports were described as too rough for girls or girls were described as not tough enough for these sports. However, several children were unable to say exactly why they thought girls should not do a particular sport. Instead, they referred to girls playing rugby as 'a bit strange', 'odd', 'weird' or, as Stuart did, they just felt it was not quite right.

Stuart : I dunno but like girls....it's a like a men and boy's game rugby, like girls like well it's not like fitting into it, it's just like a men's game and boys.

As mentioned previously, the female respondents seemed to reflect the boys' views of girls in sport, though to a lesser extent. For example, girls did admit that it would be odd or strange for a girl to take part in rugby or boxing but they did not rationalise their responses using the kinds of extreme stereotypes that the boys did. Elaine, for instance, when asked what she would say to a girl who took up rugby said "I'd say she could do it really", but asked what she would feel 'inside' without saying it out loud said "I'd say 'Oh she's a bit strange'[slight laugh]". When asked to explain her reaction, rather than suggesting for example that girls are too weak or too submissive, she suggested that there was just something not quite right about girls playing rugby:

Elaine : Because it's it's not like a normal sport for girls would do like....if I said I was starting like trampolining like they would say "Oh that's nice" and like cos it's a girls' sport and like a rugby if rugby it's not for girls it's I don't think it's for girls.

This is not to say that girls do not denigrate girls indirectly. For instance, Carol was asked why she thought her female friend went along to a rugby club.

Carol: I don't know I just think she likes going just for a bit of fun...and have a run around....and to meet other people.

Similarly, when asked if she had ever considered going along to the rugby club with her friend, Carol said she might so that she can have 'a bit of fun', stay fit and meet people. When asked if her friend enjoyed the aggressive side of rugby, Carol said that it would depend on how serious it was:

Carol: If it was really really aggressive then I don't think that she would join in but if it was just a tiny bit then I think she would.

This seems to suggest that if females want to take part in male-dominated sports such as rugby they have to downplay their motives so that they do not appear to be seriously competitive or aggressive. That is, it may be acceptable for girls to play rugby for social reasons but not because they like rough and aggressive sport. Thus, although girls may be playing rugby more, there may still be some kind of taboo against females taking part in very aggressive and physical sports. Of course, there is nothing to stop anybody taking part in sports for social reasons and to keep fit. However, if this was all that Carol's friend wanted, there are many sports which would fit those criteria.

In contrast to Carol's view is that of Deborah, who reported that a female friend was disappointed when they had to play touch rugby at primary school. Her friend, who had played rugby before, wanted to tackle properly as it was all 'part of the game'. This suggests that, at least for this girl, it was acceptable for females to engage in aggressive, contact sports and to express an enjoyment of these features as reasons for participation.

Overall, the participants in Study Four expressed general disapproval of females taking part in sports such as rugby or boxing. However, while explanations sometimes referred to females as weak, prone to injury and unaggressive, the most common rationalisation was that there was something strange or odd about females playing rugby. However, this view was frequently expressed along with a statement of 'right' i.e. that girls should be allowed to play whatever sports they want to.

Attitudes towards males in sport

Compared to attitudes towards rugby-playing females, the resistance to boys taking part in 'feminine' sports such as netball was far stronger. Kenneth said, quite openly, that he would call a boy playing netball 'a big ponce'. This derision was matched by several of the other children (mostly the boys) who would describe the boy as 'cissy', a 'poof', a 'nancy' and, significantly, 'a girl'. All these terms were seen as insulting to

boys, with being called a 'girl' the most common, and apparently most extreme, form of abuse. These terms of abuse appear to reflect a loss of 'masculinity'.

In explaining their responses, the children appeared to believe very strongly that, despite similarities, netball is a girls' sport and basketball is a boys' sport.

Elaine: ..but I think basketball and netball like the better one I think is basketball cos they're quite the same but like that's for them [boys] basketball so....I think netball they [boys] shouldn't play.

When asked why boys could not play netball when it is very similar to basketball, several of the boys resorted to stereotypic views of males and females to justify their views. These stereotypes seemed to typify females (and the sports they do) as slow, unenergetic and unchallenging, whereas boys (and boys' sports) are fast, skilful and energetic. Steve suggested that basketball is "...much more energetic as well cos like you're running with the ball all the time". Along similar lines, Pete said "...basketball played by men is much faster than the girls' netball". Adam said that he wouldn't like to play netball because he likes running around and being 'active'. Keith even suggested that netball "...isn't a real sport". Stuart though was less sure:

Stuart : I dunno it's just, it's a girls' game, and like boys like don't really fit in with it.

This, again, seems to reflect a general attitude that boys should take part in boys' sports and girls should take part in girls' sports. While some children were able to explain their answers, others just expressed a view that it did not 'seem right'. For instance, Martin felt that boys should play whatever sports they wanted to but still felt that playing netball was 'silly'.

Martin : It sounds a bit silly a boy playing netball but before I said [laughs] girls can...boys can play whatever they want to it sounds a bit silly really.

What is clear from these interviews is that, in contrast to the results from Studies One and Two, these children expressed very strong attitudes towards males and females in sport. Girls playing rugby were seen as different and odd, and boys playing netball were seen as effeminate and emasculated. The reasons given for why girls should not play rugby, or why boys should not play netball, varied but generally reflected stereotypes of females as weak, slow, unaggressive and prone to injury, and boys as aggressive, strong, energetic and able to take pain. In these respects, the attitudes of the children in this study seem very similar to traditional views of males and females in sport, such as those reported by Scruton (1992). One exception may be that the interviewees in the present study held less strong negative attitudes towards females playing football.

Social Norms and the Expression of Attitudes

It remains to be answered why children show clear evidence of stereotyping in the interview situation but not when completing the questionnaires in Studies One and Two. As previously discussed, social judgeability theory suggests that when making judgements of other people we respect social norms regarding how we make judgements. Leyens et al. (1994) suggest that people know it is unacceptable to judge others on the basis of their category membership i.e. whether they are male or female. The interview data show clear evidence that when expressing their attitudes the children were conforming to social norms such as this. In a variety of different ways, the participants expressed stereotypic attitudes, similar to traditional views of the sexes in sport, at the same time as expressing more liberal attitudes. For instance, several children predicted strong negative reactions to children in their class who had chosen to take part in sex-inappropriate activities. However, when questioned further, these children often said that they themselves would not express these negative views but that the other children in their class would. For example Martin, when asked what his reaction would be to girls playing rugby said:

Martin : Well, my opinion I wouldn't do anything I'd just go "Oh that's good" [slight laugh] but some people might think that boys are stronger than girls and that em it's too rough a game for girls.

Another striking feature was the type of reply children gave to the question "Are there any sports that you think girls/boys shouldn't do?". Almost without exception the participants replied by saying something like "No, they can do what they like" or "People can do whatever sport they want to". This seemed very much to be a statement of a principle that was not always reflected in their other attitudes. This is illustrated by Allan who stated that there were no sports that girls should not do, saying "They can do what they like" but later when asked his reaction to a girl playing rugby said:

Allan : We'd say "You've turned into a man" or something.

When asked specifically if he would say that, he reverted to his original statement and said that he himself would not. Similarly, Pete started off stating his own opinion, using the first person, and then changed to refer to what other people would think:

Pete: It's like I said before they're not built for it, they're like brought up like like they do all this knitting and all that like boys go out and work and that and they just stay at home mostly that's what people would say but I think that's a bit sexist.

Pete suffixed his view with "... that's a bit sexist" which appears to demonstrate awareness of what society might think. This could be interpreted as awareness of social norms as defined by social judgeability theory. Similarly, Pete was very clear that he did not think that girls should play rugby, but was prepared to accept that they should be given a choice:

Pete: It's really up to them if they want to get hurt and that it's up to them, but I don't think they should do it.

James said that he would not want to play netball "cos it's a girls sport", and thought that the boys in his class would laugh at their friend if he took up netball and would call him a 'girl'. However, James himself denied that he would say that, instead he would say "He can play whatever sports he wants".

Elaine was open in saying that she would think a girl playing rugby would be "a bit strange". However, she would not say this to the girl instead she would say that "...she could do it really" (rugby). Similarly, she suggested that a boy would be "a bit stupid to join" a netball club, but she would say to him "do whatever you want". Joanne had a similar reaction to the idea of a boy playing netball:

Joanne : Erm...I'd probably think it was a bit strange, cos not many boys do that but I suppose if that's what he wanted to do then he could go along and do it...it's his life.

Overall, most of the children endorsed some sort of liberal attitude towards males and females in sport. However, this view seemed to be contradicted by their views that boys playing netball, and girls playing rugby or taking part in boxing, were a bit strange or odd. The boys seemed to have no inhibitions about teasing another boy for playing netball and calling him a 'girl' or some similar insult. However, they knew this to be unfair. Boys seemed very reluctant to deride a girl playing rugby - perhaps this is the social norm which is most pertinent in current times. The feminist movement has created awareness of sexism and its effects on females and the children know it is wrong to deny females opportunities or to espouse the view that they should not take part in certain activities. However, it still seems acceptable to laugh at boys taking part in activities traditionally associated with females.

Discussion

In summary, the results from Study Four demonstrate that children do hold strong attitudes towards males and females in sport. However, there is also evidence that the children were aware of their attitudes and also of social norms concerning the rights of individuals to participate in whichever sports they choose.

One of the aims of Study Four was to illuminate the results of the previous three studies which found no evidence of strong attitudes towards males and females in sport. The results of Study Four provide strong support for social judgeability theory (Leyens et al., 1992) and the suggestion that through an appreciation of social norms, the children in Studies One and Two were able to determine whether it was appropriate or not to express these attitudes. Social judgeability theory suggests that confidence, itself influenced by social norms, influences whether an individual will withhold or vary their judgements across different situations.

However, it remains unanswered why the children expressed stereotypic attitudes in the interview study but not in the questionnaire studies. Social judgeability theory would suggest that the children in Studies One and Two were not confident enough to make judgements whereas the participants in Study Four were. If it is possible to establish which features of the methodologies influence confidence, then aside from learning about the content of stereotypic attitudes, more might be learned about when stereotypes are expressed.

One major difference between the questionnaire studies and the interview study was that the children were asked to make judgements about a named individual in the former, while in the latter they were asked to make more general judgements. Leyens et al. (1994) propose that the norm 'one should not judge *a given* individual on the basis of stereotypical information only' (emphasis added) is particularly prevalent in Western society today. This might explain why the children did not express stereotypes in Studies One and Two when they were asked to judge a person who was named. In Study Four, the children were asked in general about the sports that

boys and girls should take part in, and when considering individuals the questions were worded so that no actual person was referred to (e.g. 'Say one of the boys in your class decided to take up netball'). This may have led the children to feel much more confident about making judgements in the interviews than in the questionnaire studies.

Another feature of Studies One and Two which may have influenced confidence was the environment in which the judgements were elicited. Great care was taken to emphasise to the children that they were not taking part in an exam. Nevertheless, the methodology necessitated a situation that was very similar to an examination. The children sat in rows, were asked to concentrate on their own work, and were only allowed to complete the questionnaire after the researcher had explained the instructions. Some teachers even referred to 'examination conditions' when giving instructions to the children. If this led the children to feel that they were being evaluated in some way, then this could mean that the children were more likely to provide socially desirable answers. On the other hand, the interviews may have provided a situation for the children which did not instantly remind them of an examination. The children apparently enjoyed taking part in the interviews, mainly because they had time out of a class that they felt was boring. This may have resulted in a lowering of any perceived evaluation apprehension.

A feature of the questionnaires that may have had a crucial effect on the children's confidence in making judgements was the wording of the instructions. The questionnaires in Study One and Study Two included the following instructions:

Now, although you only know a little about Steven, I want you to imagine what kind of person you think he is. Read each of the words and phrases in this list and then indicate how well you think that word or phrase describes Steven.

The intention behind the words 'although you only know a little about Steven' was to put the children at ease. It was thought that without this reassurance the children might feel that the task was too difficult. However, with hindsight, it can be seen that drawing the children's attention to how little information they had been given could in fact have lowered their confidence in their entitlement to make judgements about the character in the story. In effect, following the presentation of the vignette, the participants were immediately reminded of the social norm that Leyens et al. suggest is so important. In order to test the predictions of social judgeability theory, Study Five was designed to test the effect of changing the wording of the instructions. This study is described in Chapter Seven.

Conclusions

The interviewees in Study Four demonstrated strong attitudes towards males and females in sport. While disapproval of girls boxing or playing rugby was evident, even stronger disapproval of boys playing netball was apparent. Justifications for these attitudes revolved around traditional stereotypes of males and females. Males were seen as fast, energetic and tough, while females were seen as slow, weak and prone to injury. Despite these stereotypic attitudes, there is clear evidence that the children were aware of social norms concerning the rights of individuals to participate in the activities that they wanted to.

There are several reasons that might explain why the children in Studies One and Two did not demonstrate stereotypic attitudes of males and females in sport. Social judgeability theory suggests that people must feel confident that they are entitled to make a judgement before they will express their attitudes. There are features of the methodology employed in Studies One and Two which may have influenced children's confidence. Study Five, described in the next chapter, was designed to test the possibility that the instructions were worded in such a way as to highlight how little information the children had about the characters they were assessing.

CHAPTER 7: THE GOLDBERG PARADIGM (MK IV): REBECCA AND DANIEL PLAY NETBALL, RUGBY AND GO SWIMMING

Introduction to Study Five

The main aim of Study Five, as with Studies One, Two, Three and Four, was to investigate children's stereotypes of males and females in sport. However, a further important aim was to test the predictions of social judgeability theory (Leyens et al., 1994) and in particular the possibility that the wording of the instructions in Studies One and Two resulted in the participants withholding their stereotypic judgements of the target individuals.

As discussed in Chapter Six, the results of Study Four provide some evidence for the view of Yzerbyt et al. (1994) that "people respect social rules whenever they judge other persons". The instructions that were presented in Studies One and Two may have encouraged the participants to follow one such rule by emphasising that they did not have very much information about the target individual.

To test this possibility, Study Five employed a similar questionnaire to that utilised in Study Two. However, in light of the results from Study Four, the questionnaire was adapted in three main ways. Firstly, in order to test the predictions made by social judgeability theory the instructions were varied across conditions. This was an attempt to make the participants feel more or less confident in their judgements of the characters. Secondly, following the strong attitudes expressed by the participants in Study Four towards boys playing a sport traditionally associated with females, an extra level of the variable 'character sport' was added to include netball. Thirdly, the items on which the participants could rate the target individuals included fewer attributional items and more items gauging the reaction of friends and parents. Some of the attitudes which were expressed by the children in the interview study were incorporated into these items.

Manipulation of Instructions

Yzerbyt et al. (1994) carried out a study (described in Chapter One) which provides strong support for social judgeability theory. In this study, they manipulated participants' confidence in their entitlement to make judgements about a target individual. All the participants were presented with the same limited, non-diagnostic information about a target individual. All participants also took part in a dichotic listening task before judging, in the third phase of the study, the target individual on a range of dependent measures. However, in between the dichotic listening task and the final phase, half of the participants were told that during the listening task, while they were shadowing the speech in one ear, they had received individuating information about the target individual in the ear that they had not attended to. The participants were led to believe that, even though they were not consciously aware of this information, it would still have been processed to some degree. However, this was a deception; the experimenters did not, in fact, present any extra information during the dichotic listening task. It was found that those participants who had been led to believe that they had received extra information made more stereotypic judgements of the target individual than the other participants. Yzerbyt et al. suggest that these results are due to a change in the participants' feelings of confidence in their own ability or entitlement to judge the target. That is, participants who thought that they had received more information about the target felt entitled to express their opinions about the target individual. However, because these participants did not really have any extra individuating information their judgements could only be based on stereotypes.

The phrasing of the instructions in Study Five of the present research was varied in an attempt to effect a similar change in participants' feelings of confidence. As discussed in Chapter Six, the questionnaire instructions in Studies One and Two asked participants to rate the characters 'even though you do not have much information about [Rebecca/Daniel]'. It may be, that this section of the instructions, instead of reassuring participants about the task they were being asked to complete, reduced

participants' feelings of entitlement to make any judgements about the target individuals.

In Study Five, the amount of information that was presented to the participants was identical in all conditions. However, in order to change participants' feelings of confidence, the vignettes were prefixed with either 'There is a lot of information in this story so read it slowly and carefully' or 'There is not very much information in this story so read it slowly and carefully'. The former statement was intended to induce participants into feeling that they had a lot of information about the target individual and that they were, therefore, entitled to make judgements about that person. The latter statement was intended to reduce participants' feelings of confidence by emphasising how little information they actually had about the target. It was hoped that this would make the target less 'judgeable' in the eyes of the participants. After each vignette this manipulation was repeated with either 'There was not much information in this story so you may like to read it over again.' or 'There was a lot of information in this story so you may like to read it over again.'

Manipulation of Sport

Study Two investigated the evaluations of male and female characters described as playing rugby (traditionally a masculine sport) or swimming (traditionally a neutral sport). For Study Five it was decided to include a sport that is traditionally associated with females, namely, netball. The results of Study Four demonstrated quite clearly that children have especially strong views regarding boys who take part in feminine sports.

It could be argued that disguised attitude measures are only able to assess very strongly-held attitudes (Milgram et al., 1965). It is possible that, while attitudes towards females taking part in masculine sports remain strong, they are not strong enough to be assessed by a questionnaire based on the Goldberg paradigm. The effectiveness of the Goldberg paradigm as a technique to measure attitudes could then

be tested by investigating attitudes towards boys playing netball as it is clear from Study Four that children hold very strong views on this issue. If Study Five provided no evidence of negative attitudes towards boys playing netball then this would suggest that the Goldberg paradigm is not an effective technique for measuring stereotypic attitudes.

The Dependent Measures

The questionnaire in Study Two aimed to provide the children with a wide range of items on which to rate the characters. In retrospect, it can be seen that the three sections of attribution items may have appeared very similar and, therefore, confusing to the children and so it was decided to have only one section of attribution items. This section provided possible explanations of why the character played their chosen sport. To replace the other sections on attributions, a section on the reactions of the character's friends and family was included. The items in this section were developed to include some of the more common responses given by the children during the interviews. It was hoped to learn more about the Goldberg paradigm by comparing the attitudes expressed by the participants in Study Five to those expressed by the interviewees in Study Four. The section on personal items was retained as was the section of trait items. The total number of trait items was reduced, though, to facilitate the quick completion of the questionnaires. During the interviews, one of the most common descriptions of a boy playing netball was 'sissy' and so it was decided to include this term in the list of traits. To maintain a balance between positive and negative traits for males and females, 'butch' was also included.

Given the results of Study Four it may seem to be a retrograde step to design another study based on the Goldberg paradigm. However, there are several reasons why such a study would be both useful and informative. Despite the interesting data elicited from the interviews in Study Four, several of the participants demonstrated difficulty with the open-ended response format. That is, while some children found it very easy to express their opinions, others struggled to verbalise their views. This may be due

to the linguistic ability of each participant or their ability to reflect on the attitudes that they hold. Questionnaires can help overcome these problems by presenting items which the participants can endorse to varying degrees, but which the participants do not have to generate themselves. The attitudes of children with extensive vocabularies are, therefore, not as over-represented as they might be in an interview-based study. Another advantage of rating scales is that the attitudes of a large number of participants can be recorded and directly compared.

It must be remembered that if participants are aware of the nature of a study then they may answer in socially desirable ways. A questionnaire based on the Goldberg Paradigm is, therefore, useful in eliciting attitudes without the problem of reactivity. Greenwald and Banaji (1995) suggest very strongly that disguised measures are not just useful but essential. This is because stereotypes are, by their very nature, not always introspectively known to an individual. Direct measures, such as interviews, that rely on introspection are, therefore, limited.

A further reason for choosing to employ the Goldberg paradigm was that it provided the opportunity to directly test the predictions of social judgeability theory. While interviews can provide invaluable data concerning children's attitudes they are unable to test specific hypotheses in the way that an experimental study can.

Swim et al. (1989) report that over 103 studies based on the Goldberg paradigm were published between 1968 and 1985. It is important to learn as much as possible about the use of this paradigm in stereotyping research. If the wording of the instructions is crucial to the expression of stereotypes by participants then this needs to be established. It was hoped that the results of Study Five would generate useful information regarding when and under what conditions children express stereotypic attitudes of males and females in sport.

In summary, the main aim of Study Five was to investigate children's attitudes towards males and females in sport. A further aim was to test the predictions of social judgeability theory by manipulating the instructions given to the participants.

Method

Design

This study comprised a $2 \times 3 \times 2 \times 2$ factorial design varying the sex of the character in the story (Daniel or Rebecca), the sport of the character in the story (rugby, swimming or netball), the sex of the participant (male or female) and the wording of the instructions ('lots of information' or 'not very much information').

Questionnaire Design

Vignette Design

The vignette used in Study Five was based on that used in Study Two. In order for the manipulation of confidence to be successful it was necessary that both types of instructions were credible. That is, it was important that the children who were told that there was 'a lot of information in the story', actually believed that there really was a large amount of information. Similarly, those children who had their attention drawn to how little information there was in the story, should have been led to believe that they had not been given much information about the character. For these reasons, it was decided to add a couple of extra paragraphs to the vignette to give the appearance of a lot of information. However, care was taken to ensure that these extra paragraphs contained no diagnostic information about the character and in this way maintain consistency with the 'little information' conditions. The complete vignette was as follows:

Rebecca Smith, aged 15, a pupil at West Holden Comprehensive School, has just been selected to take part in a summer school organised by Cambridgeshire Schools Swimming Association. Along with 45 other girls from around the county, Rebecca will spend three weeks at Foresham Hall. Each day the swimmers will practise the skills and techniques essential to swimming, and follow a challenging fitness training programme. Rebecca said "I am really looking forward to it. Three weeks

of hard training will be tough, but I am sure that I will learn a lot from it." Rebecca was unsure how far she would go in the sport. "Well it's a bit early to say. I would like to swim for the county team one day. Some of the girls at the summer school will be very good, and I will need to work hard to keep up with them."

One of the coaches organising the summer school said "Rebecca is very keen and eager to learn, and a great example to the rest of the girls. If you want to succeed at this sport you have to be prepared to work hard. Of course we want everybody to have fun too, and I am sure that Rebecca will enjoy her three weeks with us."

One of the PE teachers at Rebecca's school said "I am really pleased for Rebecca. She has worked really hard all year and going to the summer school is a great reward. She is a fine example to the rest of the school." Rebecca's geography teacher joked "I wish she worked as hard at her schoolwork as she does at swimming!"

Instructions To Participants

The instructions asked participants to read the story and told them that they would later be asked to describe the main character. In half of the questionnaires the vignette was prefixed with 'There is a lot of information in this story so read it slowly and carefully' and followed by 'There was a lot of information in this story so you may like to read it over again.' In the other questionnaires the instructions included the line 'There is not very much information in this story so read it slowly and carefully'. This manipulation was repeated after the vignette with 'There was not much information in this story so you may like to read it over again.' These statements were **emboldened and underlined** to make them prominent.

Manipulation Check

To check that the manipulation of the instructions did influence participants' feelings of confidence in the predicted ways, immediately after reading the vignette, the participants were asked to indicate how much they knew about the character in the story on a ten-point scale ranging from *nothing* to *everything*. A pilot study was carried out with 73 Year 7 children (12 and 13-year-olds) to ensure that the manipulation of confidence was effective. For the 'lots of information' conditions the mean score for knowledge of the character was 5.56 (SD 1.65) and for the 'not much information' conditions the mean score was 4.86 (SD 1.78).

While the results from the pilot study suggest that the manipulation was effective, it may be the case that participants would have felt more confident in their judgements after they had completed the questionnaire. That is, once they had answered all the questions about the character they may have felt that they actually did know quite a bit about the person. This may involve some kind of post-hoc rationalisation; a participant may feel that if they were able to answer all those questions about the character then they must have known quite a bit about them. However, after reading just the vignette, participants may still be unsure about how much they know. For this reason it was decided to include the manipulation check at the end of the questionnaire as well.

Item Design

In total there were four sections of items. As with the previous questionnaires, the first section included trait items which were rated on five-point Likert-type scales ranging from *very poor description* to *very good description*. The instructions for the trait items asked participants to indicate how well they thought each word described the character. No mention was made of how much information the participants had been given about the character. However, an example of how to complete the rating scales was presented for the word 'happy'. Participants were told that there was

nothing in the story to tell them if the character was happy or not, but that they could guess from the picture they had built up of the character.

The full list of trait items is shown in Table 7.1 along with possible classifications in terms of desirability (positive or negative) and of whether the trait is perceived as being a 'masculine' or 'feminine' adjective. 'Sissy' was included because of its frequent use by boys in the interviews to describe a boy playing netball. 'Butch', while used less in the interviews to describe rugby-playing girls, was included to ensure a rough balance between positive and negative traits for males and females.

	Desirability (positive or negative)	Masculine or feminine
Competitive	Positive	Masculine
Strange	Negative	Neutral
Tough	Positive	Masculine
Attractive	Positive	Neutral
Friendly	Positive	Neutral
Hard working	Positive	Neutral
Masculine	Positive for male Negative for females	Masculine
Unusual	Negative	Neutral
Sissy	Negative for males	Feminine, but used to describe males
Trendy	Positive	Neutral
Aggressive	Positive	Masculine
Athletic	Positive	Masculine
Feminine	Positive for females Negative for males	Feminine
Gentle	Positive	Feminine
Rough	Positive	Masculine
Butch	Negative	Masculine, but used to describe females

Table 7.1 Trait Items with Possible Classifications on Desirability and Masculine/Feminine Dimensions.

Section Two of the questionnaire consisted of 'evaluative' items, asking the participants how much they liked the character, how well they thought the character would do on their exams, and how likely it was that the character would fulfil their ambitions. All of these questions were answered using five point Likert-type scales. These items are shown in Table 7.2 along with the labels for the extreme points of the rating scales.

	Point 1 label	Point 5 label
How much do you like Daniel?	not at all	a lot
Would you like him as a friend?	not at all	very much
How well do you think he will do in his exams?	will fail everything	will pass everything
Do you think Daniel should spend more time or less time on his sport?	a lot less time	a lot more time
How likely do you think it is that one day Daniel will... ...swim for the county team? ...get married? ...go to college ...have children? ...become a famous sports star? ...become a coach? ...represent his country at swimming? ...give up swimming?	very unlikely	very likely

Note: For 'masculine' and 'feminine' sport conditions 'swimming' was replaced with 'rugby' or 'netball' respectively. For conditions with female characters the name and all pronouns were changed.

Table 7.2 Evaluative Items with Labels for Extreme Points of Likert-type Scales

The third section of items consisted of a variety of attributional explanations of why the character participated in their chosen sport. These items are listed in Table 7.3 along with possible classifications on dimensions of stability, locus of control and desirability. Each item was answered using a five point Likert-type scale ranging from *very likely explanation* to *very unlikely explanation*.

	Locus of Control (internal or external)	Stability (Stable or unstable)	Desirability (positive or negative)
He does it to keep fit	internal	unstable	positive
He enjoys competitions	internal	stable	positive
He enjoys physical exercise	internal	stable	positive
He loves all sports	internal	stable	positive
He enjoys being different	internal	stable	negative
The training improves his good looks	internal	unstable	negative
He wasn't good enough to get into any other sports team	external	unstable	negative
Being in the rugby team means he is popular at school	external	unstable	negative
He loves rugby	internal	stable	positive
A lot of his friends play a lot of rugby so he wants to join in	external	unstable	negative

Note: For 'masculine' and 'feminine' sport conditions 'swimming' was replaced with 'rugby' or 'netball' respectively. For conditions with female characters the name and all pronouns were changed.

Table 7.3 Attributional Items and Possible Classifications on Stability, Locus of Control and Desirability Dimensions

The fourth section of the questionnaire asked participants to consider the reactions of the character's friends and classmates to the character's sports participation. The items, shown in Table 7.4, were designed to include both positive and negative reactions. Some of the items (e.g. 'the boys in his/her class call him/her names' and 'his/her friends want him/her to try a different sport') were included to provide direct comparisons with the attitudes expressed during the interviews. Each item was scored on a five point Likert-type scale ranging from *very unlikely reaction* to *very likely reaction*.

	Desirability of Reaction (positive or negative)
His friends are jealous of his success	negative
His parents are proud of him	positive
The boys in his class call him names	negative
His parents wish he would do more schoolwork and less rugby	negative
Some people laugh at Daniel behind his back	negative
His friends encourage him	positive
The girls in his class call him names	negative
His friends think rugby is boring	negative
His friends want him to try a different sport	negative

Note: For 'masculine' and 'feminine' sport conditions 'swimming' was replaced with 'rugby' or 'netball' respectively. For conditions with female characters the name and all pronouns were changed.

Table 7.4 Items for Reactions of Friends and Family with Possible Classification on Desirability Dimension.

Following the section on friends' reactions, the participants were asked again to rate how much they knew about the character on a ten point scale ranging from *nothing* to *everything*. The pilot questionnaire included an open-ended section where participants were invited to write down any comments or ideas that they had about the character in the story. As this section did not elicit any useful data it was omitted from the final questionnaire.

The final section of the questionnaire recorded each participant's age, sex and, for administrative purposes, their registration class. It was decided not to ask the participants about their own sports participation as this had caused some problems in the administration of the previous questionnaires.

In total there were twelve different questionnaires reflecting all the possible combinations arising from varying the sex of the character in the story, the sport they played, and the type of instructions given to participants. The items in each

questionnaire were exactly the same except for the changes required by the factors being manipulated.

Participants

This questionnaire was presented to 236 children in Year 8 (mean age 12.8 years, SD 0.67) of a comprehensive school in the North East of England.

Procedure

While the pilot study for Study Five was carried out with pupils from the same school that had been involved with Studies Two, Three and Four, the data for Study Five were collected from a different comprehensive school. Because the participants in the other studies were effectively anonymous, it would not have been possible to ensure that the participants had not been involved in the research previously. It was, therefore, necessary to collect data from children at a different school.

The procedure for Study Five was different to that utilised in Studies One and Two. In Study Five the questionnaires were distributed to all participants simultaneously, rather than during classes spread over a period of days. All of the Year 8 pupils attended a class on personal and social development at the same time. The teacher with whom the researcher was liaising felt that this was the most appropriate time for distributing the questionnaires rather than during academic subject lessons. Distributing the questionnaires simultaneously ensured that the children could not be aware of the study before completing the questionnaire. In the previous studies it was not possible to be certain that the children had not discussed the questionnaire with children who then went on to complete the questionnaire at some later time (though the teachers suggested it was highly unlikely).

However, while distributing the questionnaires simultaneously solved one problem, this meant that the researcher was not personally able to distribute the questionnaires to the participants. The liaison teacher, however, felt that this would not be a problem providing that each class teacher was given explicit instructions detailing the procedure for the completion of the questionnaires.

On the day that the data were collected, the researcher supplied each teacher with a bundle of questionnaires (with equal number of each type) and a set of instructions. The instructions described the background to the research and explained why the protocol for the distribution of the questionnaires had to be followed rigorously. A 'script' was also included for the teachers to follow when explaining to the children what they should do (see Figure 7.1). This script was based on the verbal instructions which the researcher gave in Studies One and Two, and asked that the teachers emphasise to the children that they were not doing a test but that they should work alone. The verbal instructions were kept to a minimum to avoid significant variation between classes, and instead emphasised that the children should take great care to read the written instructions very carefully.

The teachers were asked not to help those children who had difficulties understanding the questionnaire, such as those with poor reading skills. It was explained that it was important not to influence the children in any way by aiding their interpretations. However, the teachers were asked to mark (discreetly) the front of any questionnaires which they think were completed by children who, in their opinion, had not fully understood the task required of them, or children who they had felt compelled to assist (and who, therefore, had not completed the questionnaire by themselves).

The questionnaires were returned to the researcher by the class teachers. The researcher answered any queries that the class teachers had and also gained feedback on the distribution and completion of the questionnaires. With no apparent problems, the teachers were thanked for their time and assured that they would receive feedback in the form of a written summary of results.

"What we are going to do this morning has nothing to do with PSE. We are using this period to fill in a questionnaire - it's part of a survey which the school is helping with"

"In this questionnaire you will be asked to read a story and then answer some questions about the main character"

"It is important to understand that this is not a test. It is important, though, that you all work alone. Do not look at other people's work and do not talk while you are filling in the questionnaire - that is why you are sitting as if it was a test"

"It is also important to understand that there are no right or wrong answers. It is your own ideas that are important. If there is anything that you do not understand, like a word that you don't know, it doesn't matter. You can just leave it out and go onto the next bit. However, if there is a word that you don't know, to show that you haven't missed it by accident, place a cross next to it and go onto the next one"

"All the instructions explaining what you should do are at the beginning of each section of the questionnaire - do not rush in and start answering the questions until you have read the instructions. It is not a race - read all the instructions carefully. It is sometimes tempting just to go straight ahead and answer the questions - for this questionnaire though you must take your time, and read everything carefully"

"You may now turn over and start - REMEMBER read everything carefully!"

Figure 7.1 Teachers' Verbal Instructions to Children

Results

As in the previous questionnaire studies, Study Five produced several results that suggested that the children did understand the questionnaire and the task that they were asked to do, and that they did not fill the questionnaire in randomly. There were also several results which support the findings from the interview study. However, there was no evidence that the characters described as doing 'inappropriate' sports for their sex were denigrated in a consistent way. Manipulating the wording of the instructions did result in different levels of reported confidence but did not seem to result in any more stereotypic descriptions of the characters.

There is clear evidence that the children did understand the questionnaire and that they completed it attentively. As the data in table 7.5 show, the children expressed clear stereotypes of the various sports. Rugby players were seen as more aggressive, $F(2,203)=11.92, p<0.05$, more rough, $F(2,210)=31.59, p<0.05$, and less gentle, $F(2,209)=8.25, p<0.05$, than the swimmers and the netball players. However, while the swimmers were perceived as the least aggressive of the sports participants, they were rated as the most athletic, $F(2,212)=4.99, p<0.05$. These sport stereotypes are consistent with the results from Studies One, Two and Three.

	Rugby	Swimming	Netball
Aggressive	2.70 (1.31) n=85	1.80 (0.97) n=78	2.43 (1.29) n=76
Athletic	4.64 (0.74) n=85	4.91 (0.29) n=78	4.57 (0.91) n=76
Gentle	2.16 (1.17) n=85	2.61 (1.14) n=78	2.91 (1.09) n=76
Rough	3.32 (1.36) n=85	2.04 (0.99) n=78	2.14 (0.93) n=76

Table 7.5 Mean scores, Standard Deviations and Number of participants for Aggressive, Athletic, Gentle and Rough by Sport

Results showing sex of participant and sex of character interactions also indicate that the children understood the questionnaire and its instructions. For instance, as Table 7.6 shows, male participants rated the female characters as more attractive and female participants rated the male characters as more attractive, $F(1,197)=6.02$, $p<0.05$. Interestingly, participants rated the characters of their own sex as less likely to go to college, $F(1,211)=4.34$, $p<0.05$, less likely to do well on their exams, $F(1,211)=6.51$, $p<0.05$, less hardworking, $F(1,209)=5.64$, $p<0.05$, and less likely to give up their chosen sport, $F(1,211)=5.26$, $p<0.05$.

		Male participants		Female participants	
	Label for Point 5 of Scale	Male character	Female character	Male character	Female character
Attractive	Very good description	2.02 (1.08) n=61	2.36 (1.24) n=55	2.48 (1.20) n=66	2.04 (0.96) n=54
Hardworking	Very good description	3.67 (1.46) n=61	3.98 (1.15) n=55	3.94 (1.13) n=66	3.58 (1.50) n=54
How likely do you think it is that Rebecca/Daniel will go to college?	Very likely	2.65 (1.18) n=61	3.02 (1.21) n=55	3.12 (1.16) n=66	2.85 (1.09) n=54
How well do you think that Rebecca/Daniel will do on her/his exams?	Will pass everything	2.72 (0.92) n=61	3.05 (0.97) n=55	3.08 (0.95) n=66	2.81 (0.75) n=54
How likely do you think it is that Rebecca/Daniel will give up swimming/netball/rugby?	Very likely	1.40 (0.74) n=61	1.71 (1.08) n=55	1.76 (1.08) n=66	1.50 (0.93) n=54

Table 7.6 Mean scores, Standard Deviations and Number of Participants for Attractive, Hardworking, 'College', 'Exams', and 'Give up sport' by Sex of Participant and Sex of character

As in the previous questionnaire studies, a significant interaction was found between sex of character and sport of character for the item 'unusual', $F(2,208)$, $p<0.05$. As Table 7.7 shows, the female rugby player and the male netball player were rated as the most unusual of the characters. An LSD post-hoc test confirmed that the female rugby player was rated as significantly more unusual than both the male and female swimmers, and that the male netball player was rated as significantly more unusual than both the male and female swimmers. However, the male netball player was not rated as significantly more unusual than the female rugby player. As with Study Two, no significant effects or interactions were found for the item 'strange'.

	Male character	Female character
Netball	2.20 (1.14) n=40	1.44 (0.65) n=36
Swimming	1.83 (1.14) n=42	1.54 (0.74) n=36
Rugby	2.07 (1.12) n=47	2.21 (1.17) n=38

Table 7.7 Mean scores, Standard Deviations and Number of Participants for 'Unusual' by Sex of Character and Sport.

Further evidence that the children understood the questionnaire and that they completed it attentively comes from the results for the questions asking how much the participants knew about the character in the story. Table 7.8 shows the mean scores (scale ranged from 1, *nothing*, to 10, *everything*) for the two confidence measures. The wording of the instructions was found to have a significant effect when this question was asked at the beginning of the questionnaire, $t(236)=5.26$, $p<0.05$, and at

the end of the questionnaire, $t(236)=3.29, p<0.05$. For all participants there was a significant difference between the knowledge of the character at the start of the questionnaire compared to knowledge of the character at the end of the questionnaire, $t(236)=16.35, p<0.05$. These two measures were also found to be significantly correlated, $r(237)=0.622, p<0.05$. These results suggest that manipulating the wording of the instructions did have a significant influence on how much participants felt that they knew about the characters, and that after completing the whole questionnaire, regardless of the wording of the instructions, participants felt that they knew more about the character than they did after reading only the vignette.

		Wording of instructions	
		'Lots of information'	'not very much information'
How much do you know about Rebecca/Daniel?	start of questionnaire	5.59 (1.50) n=119	4.52 (1.62) n=119
	end of questionnaire	6.97 (1.64) n=119	6.26 (1.71) n=119

Table 7.8 Mean scores, Standard Deviations and Number of Participants for 'How much do you know about Rebecca/Daniel?'

Several items from the section on 'reactions from friends and family' produced results consistent with the responses given by the participants in the interview study. Table 7.9 shows the mean scores, standard deviations and number of participants for the items 'The boys in the class call him/her names', 'Some people laugh at Rebecca/Daniel behind her/his back' and 'The girls in her/his class call her/him names'. For these items, a higher score indicates a more likely reaction.

For the item 'The boys in the class call him/her names' there was a significant interaction between sex of character and sport, $F(2,212)=5.23$, $p<0.05$. An LSD post-hoc test revealed that this reaction was rated as significantly less likely for the male rugby player when compared to the female rugby player and the male netball player. Interestingly, this reaction was rated significantly more likely for the female netball player than for the male rugby player.

	Male characters			Female characters		
	netball	swimming	rugby	netball	swimming	rugby
The boys in the class call her/him names	2.75 (1.33) n=40	2.36 (1.34) n=42	1.87 (1.01) n=47	2.47 (1.08) n=36	2.31 (0.95) n=36	2.74 (1.22) n=38
Some people laugh at Rebecca/Daniel behind her/his back	3.18 (1.11) n=40	2.64 (1.16) n=42	2.23 (1.07) n=47	2.31 (1.09) n=36	2.39 (1.02) n=36	2.82 (1.18) n=38
The girls in her/his class call her/him names	3.05 (1.26) n=40	2.10 (1.01) n=42	1.94 (0.99) n=47	2.00 (1.04) n=36	2.17 (1.00) n=36	2.61 (1.05) n=38

Note: Scale ranged from 1, *very unlikely reaction*, to 5, *very likely reaction*.

Table 7.9 Mean scores, standard deviations and number of participants for 'reactions of friends' by sex of character and sport

A similar interaction between sex of character and sport was found for the item 'Some people laugh at Rebecca/Daniel behind her/his back', $F(2,212)=7.95$, $p<0.05$. Consistent with the findings from Study Four, an LSD post-hoc test revealed that this reaction was rated as most likely for the male netball player and the female rugby player. However, while this reaction was rated as most likely for the male netball player, this was not significantly different from the score for the female rugby player.

The interaction between sex of character and sport was also significant for the item 'The girls in her/his class call her/him names', $F(2,212)=11.22$, $p<0.05$. Again, the LSD post-hoc test showed that this reaction was rated as most likely for the male netball player and the female rugby player, but that the ratings for these two characters were not significantly different.

These results are consistent with the views expressed by the participants during the interviews. Those people described as taking part in sports not usually associated with their sex are subject to a degree of derision in the form of name-calling and surreptitious ridicule. However, there were no other results which suggest that the characters taking part in 'inappropriate' sports were denigrated in a consistent way.

As previously mentioned, manipulating the wording of the instructions was effective in influencing the amount of information that the participants felt they had about the characters. However, varying the instructions did not have any consistent effects on judgements of the characters. There is no evidence to suggest that telling participants that they had received a lot of information resulted in more extreme or stereotypic judgements of the characters. One possible exception to this comes from the results for the item 'masculine'. The results for this item (see Table 7.10) did demonstrate a significant interaction between sex of character, sport and wording of instructions, $F(2,194)=6.13$, $p<0.05$.

Figure 7.2 and Figure 7.3 illustrate the results for 'masculine'. In the 'lots of information' conditions, all the female characters were rated in similar ways to the male netball player. However, an LSD post-hoc analysis shows that the male rugby player and the male swimmer were rated as significantly more masculine than the other characters. Following the predictions of social judgeability theory, it might be suggested that when participants feel confident in making a judgement they denigrate the male netball player by describing him as significantly less masculine than the male swimmer and male rugby player. However, the results from the 'not very much information' conditions do not bear out such an analysis. For instance, it is not clear

why the female rugby player was rated as slightly more masculine than the male rugby player, but the male netball player was rated as significantly more masculine than the female netball player. Varying the wording of the instructions does not seem to have had a consistent effect on participants' ratings for the item 'masculine'.

	Wording of instructions			
	lots of information		not very much information	
	male character	female character	male character	female character
Netball	2.59 (0.87) n=20	2.06 (0.97) n=18	3.61 (0.98) n=20	2.06 (0.97) n=18
Swimming	3.50 (1.26) n=22	2.47 (1.30) n=17	3.11 (1.45) n=20	1.88 (0.93) n=19
Rugby	3.76 (1.18) n=23	2.44 (1.29) n=20	2.95 (1.50) n=24	3.17 (1.15) n=18

Note: Rating scale ranged from 1, *very poor description*, to 5, *very good description*.

Table 7.10 Means, standard deviations and numbers of participants for 'Masculine' by Sex of Character, Sport and Wording of Instructions

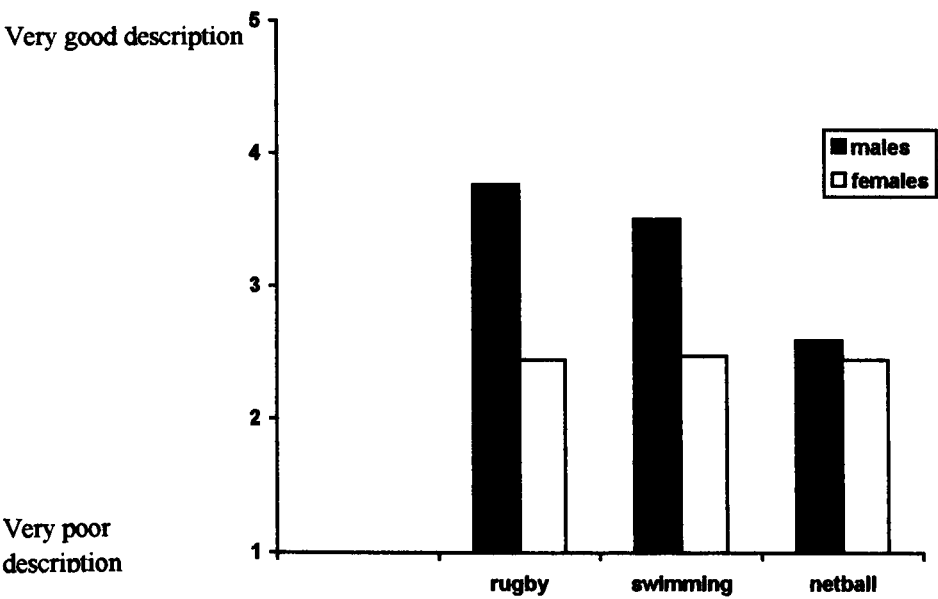


Figure 7.2 Mean ratings for 'Masculine' by Sex of Character and Sport in 'lots of information' conditions.

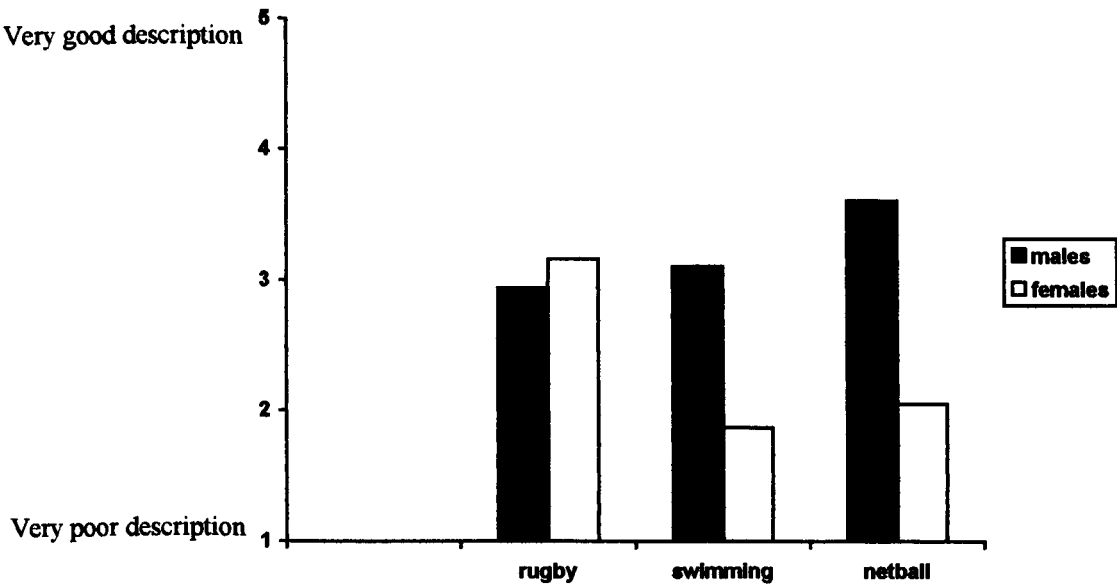


Figure 7.3 Mean ratings for 'Masculine' by Sex of Character and Sport in 'not very much information' conditions.

The possibility that the results for 'masculine' were spurious is supported by the lack of any similar interactions for any of the other items. In particular, the complex

interactions elicited for 'masculine' were in stark contrast to the results for 'feminine'. As expected, the female characters ($M=3.19$, $SD\ 1.41$) were rated as significantly more feminine than the male characters ($M=1.79$, $SD\ 1.17$), $F(1,197)=64.6$, $p<0.05$, but there were no interactions with any of the other factors.

Results that could be interpreted as denigration of characters who take part in 'inappropriate' sports come from the item 'How likely do you think it is that one day Rebecca/Daniel will have children?' (see Table 7.11). The interaction between sport and sex of character was significant, $F(2,211)=4.56$, $p<0.05$, and an LSD post-hoc analysis confirmed that the male netball player and the female rugby player were rated as less likely than the other characters to have children.

	Sex of Character	
	Male	Female
Netball	2.92 (1.13) n=39	3.44 (1.11) n=36
Swimming	3.43 (1.15) n=42	3.22 (0.96) n=36
Rugby	3.49 (0.93) n=47	3.00 (0.90) n=38

Table 7.11 Mean ratings, standard deviations and numbers of participants for 'How likely do you think it is that one day Rebecca/Daniel will have children?' by Sex of Character and Sport (scale: 1=very unlikely, 5=very likely)

The mean ratings of likelihood for the reaction 'Her/his parents are proud of her/him' are shown in Table 7.12. A significant interaction was found between sex of character and sport, $F(2,212)=3.95$, $p<0.05$. An LSD post-hoc analysis revealed that this reaction was rated as considerably less likely for the female rugby player compared to all the other characters (see Figure 7.4). While this could be interpreted as denigration of a female taking part in an inappropriate sport, it is not clear why the male character described as taking part in an inappropriate sport was not denigrated in

a similar way. The results from Study Four would suggest that males are subject to stronger disapproval than females for taking part in sports considered inappropriate.

	Sex of Character	
	Male	Female
Netball	4.38 (1.03) n=40	4.64 (0.68) n=36
Swimming	4.52 (0.92) n=42	4.78 (0.48) n=36
Rugby	4.43 (0.93) n=47	3.95 (1.27) n=38

Table 7.11 Mean ratings, standard deviations and numbers of participants for ' Her/his parents are proud of her/him ' by Sex of Character and Sport (scale: 1=*very unlikely*, 5=*very likely*)

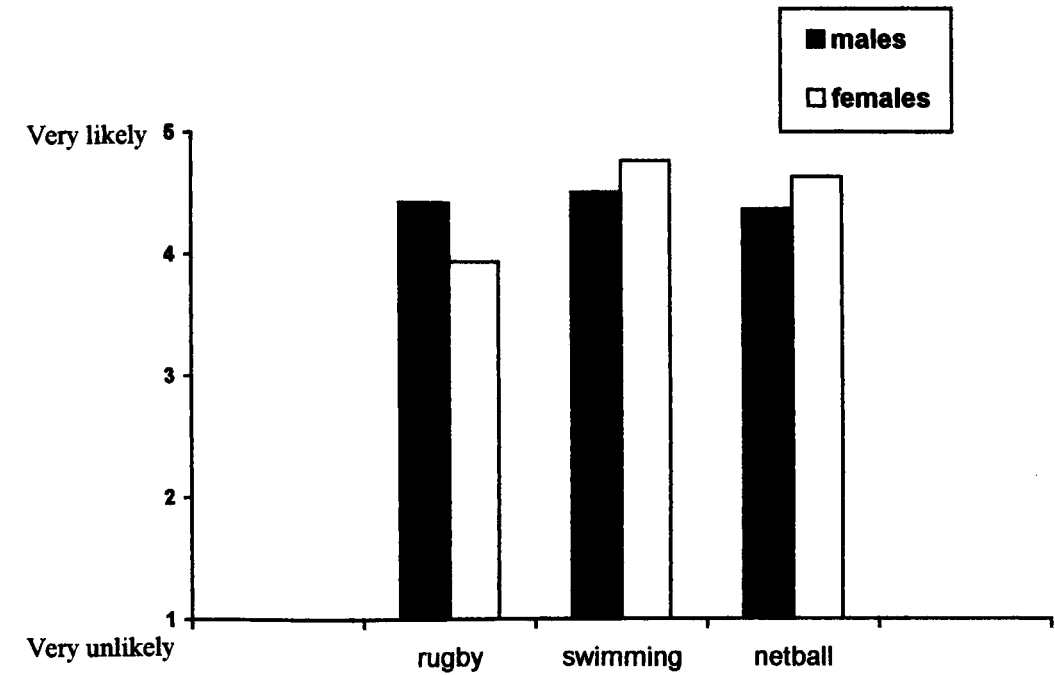


Figure 7.4 Mean ratings for 'Her/his parents are proud of her/him' by Sex of Character and Sport

Apart from the items already mentioned, no consistent interactions between sex of character and sport were found. That is, as with Studies One and Two, there was no consistent evidence to suggest that those characters who were described as taking part in 'inappropriate' sports were denigrated or derogated. While there were some notable interactions that suggest the male netball player and the female rugby player were perceived, and subsequently described, differently, the different descriptions did not appear to necessarily reflect negative evaluations. Due to the large number of conditions it is not possible to show here the results for all items in the questionnaire. However, the raw data set in SPSS for Windows format can be found in Appendix H.

The manipulation of the wording of the instructions was effective because participants who had been told that they had 'lots of information' reported that they knew significantly more about the character than those who had been told that they had 'not very much information'. The results of Studies One, Two, Three and Four suggest that the apparent lack of stereotyping in these studies was because participants were actively respecting social norms. It was hypothesised, therefore, that those participants who felt that they had a 'lots of information' would make more stereotypic judgements of the characters than those who had 'not very much information'. That is, it was predicted that participants who were confident that they were entitled to make a judgement about the character, because they had enough information to do so, would express their 'true' feelings rather than withholding them in respect of social norms. This would be reflected in significant interactions between sex of character, sport and type of instructions. In particular, it was expected that the characters described as taking part in 'inappropriate' sports would be evaluated negatively only by those participants who were told that they had been given 'a lot of information'. However, apart from the effects found for the item 'masculine', no such interactions were found to support these predictions.

It is clear, therefore, that manipulating the presentation of the information did not elicit any more gender stereotypes than in the previous studies. One possible explanation for the results in Study Five not supporting predictions concerns the nature of the dependent measures. For instance, the main dependent measure

employed by Yzerbyt et al. (1994) was the number of *don't know* answers provided by participants. These researchers found that those participants who had had their confidence manipulated so that they felt that they knew more about the target individual responded using *don't know* less often than those participants whose confidence was not manipulated. The ANOVA analysis employed in the present study compares the mean scores from the five-point Likert-type scales. This type of analysis is not able to distinguish between conditions where everybody endorses the midpoint of the scales and conditions where exactly half of the participants endorse 1 and the other half endorse 5, as in both cases the means would be 3 (though the standard deviations would be different).

It was, therefore, decided to carry out an analysis which could distinguish between these possible response patterns. The five-point Likert-type scales were converted to three-point scales of confidence. It was assumed that participants who endorsed scores of 1 or 5 were the most confident in their judgements, and so were recoded as 2. Similarly, those participants who endorsed scores of 3 were assumed to be the least confident in their judgements and were recoded as 0. The intermediate scores of 2 and 4 were recoded as 1. The individual confidence scores for the items within each section of the questionnaire were then totalled to give five overall confidence scores. The sections were 'traits', 'personal', 'likely life events', 'Why does R/D play rugby/go swimming/play netball?' and 'reactions of friends and family' (corresponding to the different sections in questionnaire). These five confidence measures were also totalled to give one overall confidence measure for the questionnaire. Pearson's correlations were then employed to investigate the relationships between the measures of 'extremity of judgements' and the ratings of how much participants felt they knew about the characters. Table 7.12 shows the results for these correlations. It should be noted that total scores were not calculated for participants with missing values. The large number of missing values for the trait items resulted in total scores for only 161 participants (out of 236 in the total sample).

The results shown in Table 7.12 provide some evidence to suggest that those participants who felt that they knew most about the characters were also those who

made more extreme judgements. This is more apparent for the second measure of how much the participants knew about the character they had read about. In particular, those participants who gave more extreme ratings for the personal items, the items relating to the reactions of friends and family and the attributional items, also reported knowing more about the character in the story. Conversely, participants who tended to endorse the midpoints of the Likert-type scales for these items also tended to report that they did not know much about the character in the story. The extent to which these correlations support the predictions of social judgeability theory will be discussed in the next section. However, it should be noted that, while some of the correlations indicate significant positive relationships, the coefficients are small.

Measures of 'Extremity of Judgements'	How much do you know about Rebecca/Daniel?	
	After vignette	After questionnaire
Trait items	r=0.004 n=172 non-sig	r=-0.003 n=171 non-sig
Personal items	r=0.112 n=235 p=0.043	r=0.165 n=235 p=0.006
Likely life events	r=0.078 n=231 non-sig	r=0.081 n=231 non-sig
Why does Rebecca/Daniel play rugby/ play netball/go swimming (attributional items)	r=0.162 n=231 p=0.007	r=0.207 n=231 p=0.001
Reactions of friends and family	r=0.075 n=236 non-sig	r=0.155 n=235 p=0.009
All sections	r=0.104 n=161 non-sig	r=0.157 n=160 p=0.024

Table 7.12 Pearson Correlation Coefficients, Number of Participants and Significance Levels for Relationships between Confidence Measures and Extremity of Judgements Measures (significant results in shaded cells)

Discussion

In summary, the results from Study Five showed that the manipulation of the wording of the instructions to the questionnaire was effective. Participants who were told that they had received 'a lot of information' felt that they knew significantly more about the character than those participants who had been told that they had received 'not very much information'. However, contrary to expectations, varying the instructions did not result in consistent negative gender stereotyping of the characters.

At first glance, it may appear that some evidence to support the predictions of social judgeability theory was found; those participants who felt that they knew more about the character after completing the questionnaire tended to be more extreme in their ratings of the characters. However, there are two problems with the results of the 'extremity of judgements' measures. Firstly, the correlations of these measures with the ratings of how much the participants felt they knew about the character were very low and were only significant for some sections of items. Correlation coefficients as low as 0.200, while possibly statistically significant, are only able to account for very small proportions of variance. This limits any explanatory value that these correlations might have. Secondly, and more importantly, it is not clear why participants who felt that they knew more about the characters should make more extreme judgements in both directions, i.e. the characters were described in stereotypic and counter-stereotypic terms. Yzerbyt et al. (1994) found that participants who felt that they knew more about the target avoided the mid-point of the rating-scales and became more extreme in their judgements, but that these judgements were in the direction of the stereotype. So, for instance, a comedian was rated as more extroverted. In the present study, it appears that participants may have tended to avoid the mid-point of the rating-scales slightly more when they felt they knew more about the character. However, their more extreme judgements did not reflect any consistent stereotypes. This problem and the weak correlations suggest that the results of Study Five should not be interpreted as evidence to support the predictions of social judgeability theory. Nevertheless, the significant but low correlations may suggest that further investigation is warranted.

It remains to explain why the manipulation should have been effective but no evidence of negative gender stereotyping was found. There are several features of this study which may have contributed to the results. The questionnaires used in Studies One and Two included in the instructions 'Even though you do not know much about the character in the story, try and imagine what they are like'. It was hypothesised that this would have reduced participants' confidence in their entitlement to make judgements, and the questionnaire in the present study was designed to test this suggestion. There are, however, other factors which may have reduced participants' confidence. For instance, in the present studies there was no real evidence that the characters in the stories were real. In the study by Yzerbyt et al. (1994) the initial information about the target individual was presented in the form of a tape-recorded interview. During this interview the target individual revealed certain personal information which allowed him to be categorised. There seems to be little doubt that when the participants in Yzerbyt et al.'s study were making subsequent judgements, they felt that they were judging a real person. It is not clear in the present study whether the children believed that the person was real or not. The children were not explicitly told that the character was real, so as not to deceive them unnecessarily. Whether they thought the person was real or not, may or may not be important, but even if the participants did think that the target was a real person, making an impression based on a short written story is not the same as making an impression based on listening to an interview. In the same way that a picture is said to paint a thousand words, participants may feel that interviews are very informative even if they contain very little individuating information.

Another difference between the present study and that of Yzerbyt et al. (1994) is that in the present study the children were asked to try and imagine what the character was like and then describe them on a variety of scales ranging from 'very poor description' to 'very good description'. On the other hand, the adult participants in Yzerbyt et al.'s study were asked to complete a questionnaire in the same way as the target individual had. This may have had the effect of reducing the feelings that participants had that they were judging someone. That is, a participant pretending to be someone else may

not feel that they are judging that person. In the present study, the children were asked directly to describe and evaluate the character. This may have been a more significant influence on the children's feelings of entitlement to judge the target individual than the wording of the instructions.

A further crucial difference between the present study and that of Yzerbyt et al. (1994) is that these authors carried out a controlled experiment, whereas the present study was effectively a survey conducted by a group of teachers. Having all of the questionnaires completed simultaneously prevented any participants being aware of the nature of the study prior to taking part. This was a possible problem with Studies One, Two and Three. However, this then meant that the researcher could not distribute all the questionnaires and ensure that all the instructions provided were identical. Nevertheless, the results from Study Five are consistent with the results from Studies One and Two in many ways, and this would suggest that distributing the questionnaires in different ways is not necessarily problematic. Also, feedback from the schoolteachers suggested that they had no problems in rigorously following the instructions that they had been given. While this may be reassuring, asking teachers to distribute a questionnaire to over 200 schoolchildren during classtime is not the same as conducting a laboratory experiment. There are many factors which cannot be controlled and which may have influenced the confidence of the participants when they were rating the characters. This may be one reason why the results from the present study are apparently at odds with those of Yzerbyt et al. (1994).

There is a more serious problem with the present study which may be able to explain the apparent lack of negative gender stereotyping. The results of Study Four suggested very strongly that males who take part in inappropriate sports are subjected to direct negative evaluation by their peers. It was, therefore, predicted, assuming that the Goldberg paradigm is an effective tool, that the male netball player in Study Five would be evaluated less favourably than the other characters. However, some written comments provided by participants who were evaluating the male netball player suggested that some participants had some understanding of the nature of the study. While a male rugby player or swimmer may not be odd or strange, a male netball

player may have appeared very peculiar to the participants in the conditions with that character. If these participants guessed the nature of the study this may have seriously influenced the way in which they completed the questionnaire. For instance, these participants may have made deliberate attempts to appear non-sexist and withheld any negative views of the male netball player.

Conclusions

The results of this study did not provide evidence of widespread negative attitudes towards males and females in 'inappropriate' sports, though the results were consistent with some of the attitudes expressed in the interview study and in previous questionnaire studies. There is tentative support for social judgeability theory, as those participants who felt they knew more about the character in the story were more confident in their judgements. However, it should be noted that while some of the correlations were significant they were fairly small.

CHAPTER EIGHT: GENERAL DISCUSSION

The main aim of the present research was to investigate children's attitudes towards males and females taking part in different sports. In particular, it was hoped to investigate the nature of negative gender stereotypes of females taking part in sports that are traditionally associated with males. By themselves, the studies that were based on the Goldberg paradigm produced little evidence to suggest that children held strong negative attitudes towards females in sport. However, taking into account the results from Study Four, the interview study, the results from the present research demonstrate the context-dependency of stereotyping and the expression of negative attitudes, and this in turn has strong implications for future research into gender and sport.

The results from the present research have been interpreted in light of social judgeability theory (Leyens et al., 1994). As discussed previously, the main tenet of this theory is that when making judgements of other people, participants will respect social norms. Of particular relevance to the present research is the social norm that suggests that it is wrong to make judgements of other people when only limited information about that person is available. Yzerbyt et al. (1994) demonstrated that their participants were prepared to make stereotypic judgements of a target individual if they had been deceived into thinking that they had received appropriate information about the target.

The results of the present research are consistent with the predictions of social judgeability theory. Specifically, the apparent lack of negative stereotyping in Studies One, Two and Three could have been due to participants' reluctance to make judgements of target individuals about whom they knew very little. This is despite the indirect methodology that was employed; even if participants are not aware of the exact nature of the study they may feel that, in general, it is wrong to judge people in certain circumstances. In contrast to the first three studies, Study Four elicited clear evidence that children hold very strong views of males and females in sport, and very

strong ideas of which sports are appropriate for each sex. Study Four also provided evidence that even when children did express negative views of males and females in sport, they did so while paying regard to social norms.

In light of these results, Study Five then attempted to identify some of the features of the methodology that might have encouraged participants to withhold their views, when previous research employing the same methodology had elicited very clear and strong stereotypes. One particular feature, the wording of the instructions to participants, was tested. However, results indicated that manipulating the wording of the instructions did not influence the extent to which participants were prepared to make stereotypic judgements of the target individuals. Although apparently contradictory to the findings of Yzerbyt et al., various differences between the respective studies were highlighted. These differences may account for the lack of stereotyping demonstrated in Study Five. In particular, the medium through which information about the target was delivered and the nature of the dependent measures may have influenced the extent to which participants felt confident enough to judge the characters.

A further crucial difference between Study Five and the study by Yzerbyt et al. concerns the nature of the topic being investigated. Yzerbyt et al. asked participants to rate a comedian on scales relating to extroversion. The present research, though, asked participants to rate males and females who were described as taking part in 'unusual' sports. The latter topic is clearly a more sensitive issue than the former. The results of Study Four demonstrated quite clearly that children are aware of sensitive topics in general, and of sexism in particular. The apparent lack of stereotyping in the studies employing the Goldberg paradigm could be because the present research was concerned with a sensitive issue that has recently attracted much media coverage. Yzerbyt et al. may have elicited stereotyping in their study because asking participants to rate a comedian as extroverted is probably not particularly contentious.

The main conclusion to be drawn from this evidence would appear to be that the Goldberg paradigm is not a robust methodology for the assessment of gender

stereotypes. This conclusion would seem to be at odds with the findings of previous research (e.g. Wolfson et al., 1989) that has utilised the Goldberg paradigm to demonstrate clear evidence of stereotyping. It could be argued that attitudes towards females in sport have weakened in recent years, and that the Goldberg paradigm, like other indirect measures of stereotyping, is only effective in assessing attitudes that are strongly held. It is still possible that, while the children expressed strong negative attitudes during the interviews, these attitudes might be less strongly held than the attitudes held by previous generations. There are several different ways in which attitudes today might be considered 'weaker' than before. For instance, Krosnick, Boninger, Chuang, Berent and Carnot (1993) suggest that attitudes can vary in terms of personal importance, extremity and accessibility, as well as along other dimensions. It is possible that the attitudes expressed in the present research, even though they appeared to be strongly held, might be less important to the children, less extreme and less accessible than the attitudes held by previous generations. Until reliable and valid measures of stereotyping are developed, it will not be possible to directly compare the attitudes expressed by children today with those expressed by participants in future studies. This makes the monitoring of stereotype and attitude change problematic.

An important issue in the present research is whether the stimulus materials in the Goldberg paradigm studies actually activated stereotypes of males and females in sport (Swim et al., 1989). It is not clear whether the participants actually accessed the stereotypes of a female rugby player or a male netball player when reading the vignettes. As previously mentioned, accessibility of an attitude is an important component of an attitude's strength (Krosnick et al., 1993). Previous research has used the time taken to report an attitude as an index of accessibility (Fazio, 1986). The children that took part in the questionnaire studies may have held strong stereotypes of males and females, and of rugby players and netball players, but their stereotypes of male netball players and female rugby players may have been less strong. During the interviews, because they had more time to consider and reflect on their views, the children might have been able to express stereotypic attitudes towards

males and females in sport. The accessibility of a stereotype may only be a confounding factor for the studies that utilised the Goldberg paradigm.

The questionnaire in Study Two was designed specifically to counter the possibility that the stimulus materials did not activate stereotypes of males and females playing sport. Unlike the vignette used in Study One, the vignette in Study Two did not include information that did not relate to the character's sport involvement. For this reason, stereotype accessibility would seem to be an inadequate explanation of the results of the present research. However, Swim et al. (1989:p423) highlight the problem incurred by making stimulus materials too direct:

...it is possible that the approaches...are neither blatant enough to command stereotypic evaluations nor subtle enough to expose any real biases that subjects may harbor.

Researchers utilising the Goldberg paradigm must, therefore, tread a fine line between directness and subtlety when designing their materials. On the one hand stimulus materials must be specific enough for particular stereotypes to be activated, while on the other, stimulus materials must be subtle enough so that the aims of the study are not revealed to the participants. Greenwald and Banaji (1995:p20), strong advocates of the use of indirect measures, suggest that "To measure individual differences in introspectively inaccessible implicit social cognition, sensitive indirect measures are needed". However, they too recognise the problems of designing such measures and describe their development as "...the most significant remaining challenge".

Even if participants are not aware of the exact nature of a study, general social norms will still influence their responses. It is interesting to note that the social norms at play in the present research must be fairly specific. Studies One, Two, Three and Four all produced evidence of some stereotyping particularly of the different sports participants. For instance, participants were prepared to characterise the rugby players as aggressive and competitive, and in Study Two the female characters were described as more studious and hard-working than the male characters. Furthermore,

while the female rugby player and male netball player were described as unusual, there was no further evidence that these characters were being evaluated negatively. It would appear, therefore, that the participants were very selective in the ways in which they expressed their views. That is, the participants were prepared to stereotype the characters on some dimensions but not on others. Thus, if social judgeability theory is correct in saying that people respect social norms when making judgements, it must account for the fact that stereotyping is not an 'all-or-nothing' phenomenon.

Self-categorisation theory (Oakes et al., 1994) offers an explanation which could account for different stereotypes being applied for different items. According to this theory, the categorisation that a perceiver will use when perceiving another person (or objects) will depend on the comparative context. In the questionnaire studies, each item determines the context in which the character is being evaluated. For example, in Study Five, when the item 'attractive' was presented, clear interactions were found between sex of participant and sex of character. It would appear that when it comes to attractiveness the salient categorisation for the participants was whether the character was male or female. Similarly, in Study Five, rugby players were described as more aggressive than the swimmers and the netball players. Within this comparative context, it would appear that the sex of the character was not particularly relevant, but that the sport of the character was.

While self-categorisation theory is able to explain why different stereotypes were used to describe the characters at different times, it does not explain why there were so few interactions between sex of character and sport of character which might indicate negative stereotyping of rugby-playing females or netball-playing males. As discussed in Chapter 1, self-categorisation theory does suggest that norms are important, but this relates to the content of stereotypes rather than their application. That is, a particular stereotype will only be applied to a target if the categorisation is in line with general societal norms. Oakes et al. (1994) term this the 'normative fit' of a category. On the other hand, Leyens et al. (1994) suggest that norms are important for the expression of stereotypes and not just in determining whether they are relevant or not. Not only must a stereotype be relevant and applicable, but certain conditions must be

met before a person will be willing to express their views or make a judgement based on that stereotype. It is interesting to note that Leyens et al. (1994) do not believe that social judgeability theory and self-categorisation theory are incompatible. There are no *a priori* reasons why norms and social rules cannot be incorporated into self-categorisation theory. After all, social judgeability theory was developed in an attempt to combine social identity theory (which was itself developed into self-categorisation theory) and the traditional cognitive theories of impression formation, such as Fiske and Neuberg's continuum model (Leyens et al., 1994).

The aim of Study Five was to isolate one particular factor, the wording of the instructions, which may have led participants to believe that they were not in a position to judge the characters. Although the manipulation of this factor did influence the amount that participants felt they knew about the characters, this did not induce negative evaluations of the characters described as taking part in 'inappropriate' sports. There are, however, many other factors that could influence participants' feelings of entitlement to judge the characters. If social judgeability theory is correct, practically any feature of the experimental setting could, in theory, influence participants' confidence. For instance, the perceived status of the researcher might exacerbate or reduce evaluation apprehension. In the present research, the researcher was a stranger to the participants. The children may have been more prepared to reveal their 'true' feelings if the researcher was someone familiar to them. Other important factors may include the extent to which the target individual is believed to be real or hypothetical, the extent to which the participants feel that they are being tested and the extent to which participants feel that they have received diagnostic information about the targets.

In short, it can be seen that many different factors will influence participants' confidence. Yzerbyt et al (1994) demonstrated this experimentally, while the present research demonstrated the large difference between the expression of stereotypes during interviews and during the questionnaire studies. Taken together, these results highlight the problems that the Goldberg paradigm has in measuring or assessing stereotypes. If a study based on the Goldberg paradigm does not elicit stereotypes, it

is not possible to conclude that these stereotypes do not exist. In fact, there appear to be many factors that might lead to participants holding strong negative views to withhold their feelings. Swim et al. (1989) report that studies based on the Goldberg paradigm provide "little evidence that subjects differentially evaluate men and women". These authors conclude that "...the complexity of the conditions under which such evaluations occur and the flexibility of social perceivers' thinking must be taken into consideration". Greenwald and Banaji (1995) dispute the findings of Swim et al. and suggest that studies based on Goldberg's original procedure demonstrate that "...Goldberg's finding is well established". The results of the present research tend to support the view of Swim et al.

It is not clear whether researchers should persevere with such a context-sensitive methodology of investigating gender stereotyping. As previously discussed, Kasof (1993) has highlighted the problems of using different names to manipulate the sex of target individuals. It is not even clear whether the Goldberg paradigm is able to assess gender stereotypes in isolation from age or social class stereotypes. Taking all these problems in combination, future researchers should consider very carefully the continued use of the Goldberg paradigm. Certainly, there appear to be many reasons why the Goldberg paradigm might demonstrate no differences in the evaluations of male and female targets even if participants hold strong stereotypes. Such a fickle methodology would seem to be of dubious value.

Such a conclusion, however, is at odds with the demands made by Greenwald and Banaji (1995) for an increase in the use of indirect measures. Indirect measures were originally designed to avoid reactivity from participants, supported by evidence from research that has demonstrated discriminatory behaviour from participants who claim not to be prejudiced. It was assumed that if participants did not know what a study was actually investigating then they would be unable to respond in socially desirable ways. The results of the present research suggest that such an assumption was naive. Even when participants are not aware of the exact topic under investigation, they are still able to respect general rules of social behaviour. If anything, the results of the present research would encourage the increased use of direct methods such as the

interview method employed in Study Four. It may be that the experimental methods suggested by Greenwald and Banaji are useful to investigate the implicit cognitive processes underlying all stereotyping. However, if researchers wish to find out more about the specific content of various stereotypes, what people feel about these stereotypes, and how these stereotypes are expressed in different situations, then direct methods may prove more fruitful. It is clear that in the present research, the interviews were more enlightening of what children believe about males and females in sport than were the questionnaire studies. Nevertheless, the questionnaire studies demonstrated quite clearly that children are aware of when it is appropriate to express their views.

If researchers do continue to investigate stereotyping using studies based on the Goldberg paradigm, then it is imperative that they are aware of factors which might influence the extent to which participants feel entitled to express their views. As norms are culture-dependent, great consideration should be given to the selection of participants. For instance, in comparison to other adult groups, undergraduate students may have very different criteria for feeling entitled to make judgements about other people. Furthermore, it is essential that published research includes full details of the conditions under which participants were asked to make their evaluations and details of the stimulus materials. In this way, it may prove possible to learn more about the conditions under which people are prepared to express their stereotypes.

In contrast to many previous studies employing the Goldberg paradigm, the present research investigated children's stereotypes. Swim et al. (1989), in their meta-analysis of studies based on the Goldberg paradigm, reported that only 6 studies out of 213 involved participants younger than college age. While the present research might suggest that the Goldberg paradigm is not an effective tool to assess stereotypes, it is conceivable that this criticism is only applicable to studies involving children. Although the results of Study Three, which employed an adult sample, were consistent with the results of the studies involving children, no evidence was gathered to establish whether the adults were also withholding their judgements. If the results of the present research are to be explained by the use of children as participants, then

it would need to be established that children are more likely than adults to withhold their stereotypes. Although very few studies have investigated children's stereotypes using the Goldberg paradigm, it would seem unlikely that this methodology would be effective for assessing adults' stereotypes but not those of children.

Although the results from the present research appear to be consistent with social judgeability theory, there are problems with any theory which invokes norms as explanatory devices. As Forsyth (1995) suggests, norms have an exceptional power to explain almost any behaviour. It is much harder, however, to predict which norms will prevail in a given situation. As Leyens et al. (1994) point out norms vary across time and across cultures and also within cultures. Because of this, norms could be used to explain the expression of stereotypic attitudes as well as the withholding of stereotypic attitudes. Researchers must be careful not to dismiss any unexpected results as being due to normative factors. In many cases, norms used to explain events have, in fact, been merely a redescription of an event rather than an explanation (Bourhis, Turner and Gagnon, 1997).

Implications for Future Research

The results of the present research highlight the problems associated with the accurate assessment of stereotypes and stereotyping. Nevertheless, future research should not baulk at attempting to assess stereotypes. As discussed in Chapter One, a variety of explanations have been put forward to explain why females do not take part in sport to the same extent as males. Understanding the nature and content of stereotypes is an essential component of any explanation of participation patterns in sport. For example, the results from Study Four suggest that there is a general belief amongst some children that females are more prone to injury than males and this is why females should not play rugby or other 'masculine' sports. It is important to understand how such a belief might influence potential or existent participants in sport. Attribution theory might suggest that a female participant who gets injured may make different and less adaptive attributions than a male participant in the same

circumstances. Similarly, explanations based on self-efficacy might suggest that an injured female participant could feel less confident in her ability to recover if she believes that females are more prone to injury than males. An appreciation of the content of gender stereotypes will advance the various explanations of sport participation patterns and aid intervention strategies (Sports Council, 1994b).

Understanding the content of stereotypes may be useful for the various explanations of female sports participation, but an understanding of the processes behind stereotyping may be even more important. It is useful to know the sort of attitudes that people hold regarding females in sport, but it is also useful to know how and when these attitudes are expressed. The results of the present research have demonstrated that people are aware of the stereotypes that they hold and they may restrict their expression depending on circumstances. For example, a teacher may not outwardly express the view that females should not play rugby in circumstances where such a view might attract criticism. However, the same teacher may express stereotypic views in different ways or different situations which, even if the views are less explicit, still have a great influence on female sport participation. Sports practitioners can benefit from an appreciation of the subtle ways in which stereotyping and the expression of stereotypic attitudes occur.

Much research has been carried out on the processes behind attitude change and the circumstances required for successful attitude change (Hogg and Vaughan, 1998). The results of the present research do not provide any suggestions of ways in which stereotypes could be changed. Nevertheless, this should remain an objective for future research. The data presented in Chapter One demonstrate that while females are taking part in sport in increasing numbers, their participation is not equivalent to that of males. Future research should, therefore, continue to monitor participation patterns in sport and also the stereotypes that people hold of males and females in sport. In this way, more can be learned about stereotype change. It would be interesting to establish whether gender stereotyping is based on similar processes to stereotyping of other groups such as ethnic minorities, people with disabilities and older people. Similarly, an understanding of gender stereotyping in sport may also

illuminate the study of gender stereotyping in other domains such as employment or education.

While the results of the present research have been interpreted in the light of social judgeability theory, it should be noted that only Study Five constituted a direct test of this theory's predictions. However, the results from Study Five were inconclusive and did not, therefore, provide any clear support for social judgeability theory. This means that further tests of social judgeability theory are required before this approach can be considered a plausible explanation of the attitudes and views expressed by the participants in the present research or of other stereotyping phenomena. While one of the promising features of social judgeability theory is that it attempts to integrate self-categorisation theory and impression formation models (these approaches to stereotyping were previously antithetical), it remains to be seen whether social judgeability theory will prosper as an approach in its own right or whether it will be subsumed by, for instance, self-categorisation theory or some other approach. Without further empirical support, the latter possibility would appear to be more likely.

Conclusions

The main aim of the present research was to investigate gender stereotypes of males and females in sport. As discussed in Chapter 1, if females are to take part in sport to the same extent as males, it is important to understand the content and nature of gender stereotypes in sport, and it is particularly important to investigate children's attitudes towards males and females in sport. The present research has shown that there are many problems in measuring and assessing stereotypes. However, the results of the present research also demonstrate the context-dependency of stereotypes. If it is accepted that stereotypes and stereotyping are important influences on self-confidence, attributions and motivational orientations, and that these in turn are important determinants of sport participation, then researchers in these areas must also take into account the context in which they carry out their

research. For instance, a spectator may be willing to express negative attitudes towards a female sports participant when watching on television, but may express very different attitudes when standing on the terraces surrounded by strangers. If the link between gender stereotyping and female sport participation is to be understood, then research must be eclectic and investigate the stereotypes held by children and adults, males and females, spectators and participants in a wide range of sporting contexts.

Katz and Braly's original study of stereotypes (Katz and Braly, 1933) attempted to describe the content of various racial and national stereotypes. Researchers soon found that the content of stereotypes changed depending on context and so the focus of research moved onto the process of stereotyping. The present research has demonstrated that the stereotyping process itself is dependent on context and this would suggest that the influence of stereotyping on female sport participation patterns will remain unclear unless research considers the factors that determine people's willingness to express their views.

REFERENCES

- Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J. and Sanford, R. N. (1950). *The Authoritarian Personality*. New York: Harper.
- Ajzen, I. & Madden, T. J. (1986). Prediction of goal-directed behaviour: Attitudes, intentions, and perceived behavioural control. *Journal of Experimental Social Psychology*, 22, 453-474.
- Allison, M. (1991). Role conflict and the female athlete: Preoccupations with little grounding. *Journal of Applied Sport Psychology*, 3, 49-60.
- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258-290.
- Asch, S. E. (1952). *Social Psychology*. Englewood Cliffs, NJ: Prentice Hall.
- Ashmore, R. D., Del Boca, F. K. & Wohlers, A. J. (1986). Gender Stereotypes. In R. D. Ashmore & F. K. Del Boca (Eds.), *The social psychology of female-male relations*. London: Academic Press.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioural change. *Psychological Review*, 84, 191-215.
- Baron, R. A. & Byrne, D. (1994). *Social Psychology: Understanding human interaction*. London: Allyn and Bacon.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42, 155-162.
- Bem, S. L. (1981). Gender schema theory: A cognitive account of sex-typing. *Psychological Review*, 88, 354-364.

- Best, J. B. (1995). *Cognitive Psychology* (4th ed.). St. Paul, MN: West Publishing Company.
- Bird, A. M. & Williams, J. M. (1980). A developmental-attributional analysis of sex role stereotypes for sport performance. *Developmental Psychology*, 16(4), 319-322.
- Blucker, J. A. & Hershberger, E. (1983). Causal attribution and the female athlete: What conclusions can we draw?. *Journal of Sport Psychology*, 5, 353-360.
- Bourhis, R. Y., Turner, J. C. & Gagnon, A. (1997). Interdependence, social identity and discrimination. In R. Spears, P. J. Oakes, N. Ellemers & S. A. Haslam (Eds.), *The Social Psychology of Stereotyping and Group Life*. Oxford: Blackwell Publishers.
- Brannon, L (1996). *Gender: Psychological perspectives*. London: Allyn Bacon.
- Brawley, L. R., Landers, D. M., Miller, L. & Kearns, K. M. (1979). Sex bias in evaluating motor performance. *Journal of Sport Psychology*, 1, 15-24.
- Breakwell, G. M. (1995). Interviewing. In G. M. Breakwell, S. Hammond & C. Fife-Schaw (Eds.), *Research Methods in Psychology*. Sage: London.
- Brewer, J., & Hunter A. (1989). *Multimethod Research: a synthesis of styles*. London: Sage Publications.
- Brewer, M. B. (1988). A dual process of impression formation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in Social Cognition*. Hillsdale, NJ: Erlbaum.
- Brustad, R. J. (1992). Integrating socialisation influences into the study of children's motivation in sport. *Journal of Sport and Exercise Psychology*, 14, 59-77.

- Bucher, C. A. & Krotee, M. L. (1998). *Management of Physical Education*. London: McGraw-Hill.
- Butt, D. S. (1987). *Psychology of sport: The behaviour, motivation, personality, and performance of athletes* (2nd ed.). New York: Van Nostrand Reinhold.
- Cherry, E. C. (1953). Some experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America*, 25, 975-979.
- Clifton, R. T. & Gill, D. L. (1994). Gender differences in self-confidence on a feminine-typed task. *Journal of Sport and Exercise Psychology*, 16, 150-162.
- Coakley, J. (1987). Sociology of sport in the United States. *International Review for the Sociology of Sport*, 22(1), 63-79.
- Craig, G. J. (1992). *Human Development* (6th ed.). London: Prentice Hall.
- Csizma, K. A., Wittig, A. F. & Schurr, K. T. (1988). Sport stereotypes and gender. *Journal of Sport and Exercise Psychology*, 10, 62-74.
- Darley, J. M. & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, 44, 20-33.
- Davis, A. (1995). The experimental method in psychology. In G. M. Breakwell, S. Hammond & C. Fife-Schaw (Eds.), *Research Methods in Psychology*. Sage: London.
- Dawes, R. M. & Smith, T. L. (1985). Attitude and opinion measurement. In G. Lindzey & E. Aronson (Eds.), *The Handbook of Social Psychology* (3rd ed.). New York: Random House.

- Deaux, K. (1984). From individual differences to social categories: Analysis of a decade's research on gender. *American Psychologist*, 39, 105-116.
- Deaux, K. & Lewis, L. L. (1984). The structure of gender stereotypes: Interrelationships among components and gender label. *Journal of Personality and Social Psychology*, 46, 991-1004.
- Denzin, N. K. (1978). *The Research Act*. New York: McGraw-Hill.
- Denzin, N. K., and Lincoln, Y. S. (1994). Entering the field of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research*. London: Sage.
- Dewar, A. & Horn, T. (1992). A critical analysis of knowledge construction in sport psychology. In T. Horn (Ed.), *Advances in sport psychology*. Champaign, IL: Human Kinetics.
- Diab, L. N. (1963a). Factors affecting studies of national stereotypes. *Journal of Social Psychology*, 59, 29-40.
- Diab, L. N. (1963b). Factors determining group stereotypes. *Journal of Social Psychology*, 61, 3-10.
- Douglas, J. & Miller, J. (1977). Record breaking women, *Science News*, 112, 172-174.
- Durkin, K. (1995). *Developmental Social Psychology: From infancy to old age*. Oxford: Blackwell.
- Dyer, C. (1995). *Beginning research in psychology: A practical guide to research methods and statistics*. Oxford: Blackwell.

- Dyer, K. (1982). *Catching up the men: Women in sport*. London: Junction Books.
- Eagly, A. H. (1987). *Sex differences in social behaviour: a social-role analysis*. Hillsdale, NJ: Erlbaum.
- Eccles, J. S. & Harold, R. D. (1991). Gender differences in sport involvement: Applying the Eccles' expectancy value model. *Journal of Applied Sport Psychology*, 3, 7-35.
- Fazio, R. H. (1986). How do attitudes guide behaviour? In R. M. Sorrentino & E. T. Higgins (Eds.), *The Handbook of Motivation and Cognition: Foundations of social behaviour*. New York: Guilford Press.
- Fielding, N. (1993). Qualitative interviewing. In N. Gilbert (Ed.), *Researching Social Life*. Sage Publications; London.
- Fife-Schaw, C. (1995). Questionnaire design. In G. M. Breakwell, S. Hammond & C. Fife-Schaw (Eds.), *Research Methods in Psychology*. Sage: London.
- Fishbein, M. & Ajzen, I. (1975). *Belief, Attitude, Intention and Behaviour*. Reading, MA: Addison-Wesley.
- Fiske, S. T. & Neuberg, S. L. (1990). A continuum model of impression formation from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology*. New York: Academic Press.
- Fiske, S. T. & Taylor, S. E. (1984). *Social Cognition*. Reading, MA: Addison-Wesley.
- Fiske, S. T. & Taylor, S. E. (1991). *Social Cognition* (2nd ed.). New York: McGraw-Hill.

- Foddy, W. (1993). *Constructing Questions for Interviews and Questionnaires: theory and practice in social research*. Cambridge, UK: Cambridge University Press.
- Fontana, A. & Frey, J. H. (1994). Interviewing: The art of science. In N. K. Denzin and Y. S. Lincoln (Eds.), *Handbook of Qualitative Research*. London: Sage.
- Forsyth, D. R. (1995). Norms. In *The Blackwell Encyclopaedia of Social Psychology*. Oxford: Blackwell.
- Gaertner, S. L. & Bickman, L. (1971). Effects of race on the elicitation of helping behaviour: The wrong number technique. *Journal of Personality and Social Psychology*, 20, 218-222.
- Gilbert, G. M. (1951). Stereotype persistence and change among college students. *Journal of Abnormal and Social Psychology*, 46, 245-254.
- Gill, D. L. (1992). Gender and sport behaviour. In T. Horn (Ed.), *Advances in sport psychology*. Champaign, IL: Human Kinetics.
- Gill, D. L. (1995). Gender issues: a social-educational perspective. In S. M. Murphy (Ed) *Sport Psychology Interventions*. Champaign, IL: Human Kinetics.
- Gill, D. L. & Deeter, T. E. (1988). Development of the sport orientation questionnaire. *Research Quarterly for Exercise and Sport*, 59, 191-202.
- Goldberg, P. A. (1968). Are some women prejudiced against women? *Transaction*, 5, 28-30.
- Golombok, S. & Fivush, R. (1994). *Gender Development*. Cambridge: Cambridge University Press.

- Gould, D. (1988). Sport psychology: Future research in youth sport research. In F. L. Smoll, R. A. Magill & M. J. Ash (Eds.), *Children in Sport* (3rd ed.). Champaign, IL: Human Kinetics.
- Gould, D. & Petlichkoff, L. (1988). Participation motivation and attrition in young athletes. In F. L. Smoll, R. A. Magill & M. J. Ash (Eds.), *Children in Sport* (3rd ed.). Champaign, IL: Human Kinetics.
- Greendorfer, S. L. (1979). Differences in childhood socialisation influences of women involved in sport and women not involved in sport. In M. L. Krotee (Ed.), *The Dimensions of Sport Sociology*. Champaign, IL: Leisure Press.
- Greendorfer, S. L. (1992a). Sport socialisation. In T. Horn (Ed.), *Advances in sport psychology*. Champaign, IL: Human Kinetics.
- Greendorfer, S. L. (1992b). Gender role stereotypes and early childhood socialisation. In G. L. Cohen (Ed.), *Women and Sport: Issues and Controversies*. New York: Sage.
- Greendorfer, S. L. & Lewko, J. H. (1978). Role of family members in sport socialisation of children. *Research Quarterly*, 49, 146-152.
- Greenwald, A. G. & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem and stereotypes. *Psychological Review*, 102, 4-27.
- Hall, M. A. (1981). *Sport, sex roles and sex identity*. (The CRIAW). Ottawa: The Canadian Research Institute of the Advancement of Women.
- Hall, M. A. (1988). The discourse of gender and sport: From femininity to feminism, *Sociology of Sport Journal*, 5, 330-340.

- Hamilton, D. L. & Trier, T. K. (1986). Stereotypes and stereotyping: An overview of the cognitive approach. In J. F. Dovidio and S. L. Gaertner (Eds.), *Prejudice, discrimination and racism*. New York: Academic Press.
- Hamilton, D. L., Stroessner, S. J. & Driscoll, D. M. (1994). Social cognition and the study of stereotyping. In P. G. Devine, D. L. Hamilton & T. M. Ostrom (Eds.), *Social Cognition: Impact on Social Psychology*. San Diego, CA: Academic Press.
- Hargreaves, J. (1994). *Sporting females: critical issues in the history and sociology of women's sports*. London: Routledge.
- Harter, S. (1978). Effectance motivation reconsidered: Towards a developmental model. *Human Development*, 21, 34-64.
- Hayes, N. (1993). *Principles of Social Psychology*. Hove, UK: Lawrence Erlbaum.
- Hendry, L. (1993). *Young people's leisure and lifestyles*. Edinburgh: Scottish Sports Council.
- Hoffman, C. & Hurst, N. (1990). Gender stereotypes: Perception or rationalisation? *Journal of Personality and Social Psychology*, 58, 197-208.
- Hogg, M. A. & Vaughan, G. M. (1995). *Social Psychology*. London: Prentice Hall.
- Horn, T. S. (1987). The influence of teacher-coach behaviour on the psychological development of children. In D. Gould & M. Weiss (Eds.), *Advances in pediatric sport sciences*, 2. Champaign, IL: Human Kinetics.
- Houlihan, B. (1997). *Sport, Policy and Politics: A comparative analysis*. London: Routledge.

- Howell, D. C. (1992). *Statistical Methods for Psychology* (3rd ed.). Belmont, CA: Duxbury Press.
- Huberman, A. M. & Miles, M. B. (1994). Data management and analysis methods. In N. K. Denzin and Y. S. Lincoln (Eds.), *Handbook of Qualitative Research*. London: Sage.
- Hyde, J. S. (1990). Meta-analysis and the psychology of gender differences. *Signs: Journal of women in culture and society*, 16(1), 55-73.
- Hyman, H. (1954). *Interviewing in Social Research*. Chicago: University of Chicago Press.
- Jackson, S. & Marsh, H. (1986). Athletic or anti-social?: The female sport experience. *Journal of Sport Psychology*, 8, 198-211.
- Judd, C. M. & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, 100, 109-128.
- Kasof, J. (1993). Sex bias in the naming of stimulus persons. *Psychological Bulletin*, 113(1), 140-163.
- Katz, D. and Braly, K. (1933). Racial stereotypes of one hundred college students. *Journal of Abnormal and Social Psychology*, 28, 280-290.
- Katz, D. and Braly, K. (1935). Racial prejudice and racial stereotypes. *Journal of Abnormal and Social Psychology*, 30, 175-193.
- Krosnick, J. A., Boninger, D. S., Chuang, Y. C., Berent, M. K. and Carnot, C. G. (1993). Attitude strength: One construct or many related constructs? *Journal of Personality and Social Psychology*, 65, 1132-1151.

- Landers, D. L. (1983). Whatever happened to theory testing in sport psychology? *Journal of Sport Psychology*, 5, 135-151.
- LaPière, R. T. (1934). Attitudes versus actions. *Social Forces*, 13, 230-237.
- Lee, D. & Newby, H. (1983). *The Problem of Sociology*. London: Routledge.
- Lenney, E. (1977). Women's self-confidence in achievement settings. *Psychological Bulletin*, 84, 1-13.
- Lewko, J. H. & Ewing, M. E. (1981). Sex differences and parental influence in the sport involvement of children. *Journal of Sport Psychology*, 2, 62-68.
- Lewko, J. H. & Greendorfer, S. L. (1988). Family influences in sport socialisation of children and adolescents. In F. L. Smoll, R. A. Magill & M. J. Ash (Eds.), *Children in Sport* (3rd ed.). Champaign, IL: Human Kinetics.
- Leyens, J-P., Yzerbyt, V. & Schadron, G. (1992). The social judgeability approach to stereotypes. *European Review of Social Psychology*, 3, 91-120.
- Leyens, J-P., Yzerbyt, V. & Schadron, G. (1994) *Stereotypes and Social Cognition*. London: Sage.
- Likert, R. A. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 140, 44-53.
- Lippmann, W. (1992). *Public Opinion*. New York: Harcourt and Brace.
- Lirgg, C. D. (1991). Gender differences in self-confidence in physical activity: A meta-analysis of recent studies. *Journal of Sport and Exercise Psychology*, 13, 294-310.

- Lirgg, C. D. (1992). Girls and women, sport, and self-confidence. *Quest*, 44, 158-178.
- Locksley, A., Borgida, E., Brekke, N. C. & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology*, 39, 821-831.
- Maccoby, E. E. & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.
- Malinowski, B. (1989). *A diary in the strict sense of the term*. Stanford, CA: Stanford University Press.
- Martin, C. L. (1989). Children's use of gender-related information in making social judgements. *Developmental Psychology*, 25, 80-88.
- Matteo, S. (1986). The effect of sex and gender-schematic processing on sport participation. *Sex Roles*, 15, 417-432.
- McAuley, E. (1992). Self-referent thought in sport and physical activity. In T. Horn (Ed.), *Advances in sport psychology*. Champaign, IL: Human Kinetics.
- McKenna, R. H. (1976). Good samaritanism in rural and urban settings: A non-reactive comparison of helping behaviour of clergy and control subjects. *Representative Research in Social Psychology*, 7, 58-65.
- Metheney, E. (1965). *Connotations of movement in sport and dance*. Dubuque, IA: William C. Brown.
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative Data Analysis: An expanded sourcebook* (2nd ed.). London: Sage.

- Milgram, S., Mann, L. & Harter, S. (1965). The lost-letter technique: a tool of social science research. *Public Opinion Quarterly*, 29, 437-438.
- Miller, A. G. (1982). *In the eye of the beholder: Contemporary issues in stereotyping*. New York: Praeger.
- Miller, G. A. (1962). *Psychology: The science of mental life*. London: Penguin Books.
- Murray, M. & Matheson, H. (1992). Competition: Perceived barriers to success. In G. L. Cohen (Ed.), *Women and Sport: Issues and Controversies*. New York: Sage.
- Nicholls, J. G. (1984). Achievement motivation: conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91, 328-346.
- Oakes, P. J., Haslam, S. A. & Turner, J. C. (1994). *Stereotyping and Social Reality*. Oxford: Blackwell.
- Oakes, P. J. & Reynolds, K. J. (1997). Asking the accuracy question: Is measurement the answer?. In R. Spears, P. J. Oakes, N. Ellemers & S. A. Haslam (Eds.), *The Social Psychology of Stereotyping and Group Life*. Oxford: Blackwell Publishers.
- Office for Standards in Education (1996). *The Gender Divide: Performance differences between boys and girls at school*. HMSO.
- Osgood C. E., Suci G. J., & Tannenbaum P. H. (1957). *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.

- Parkhouse, B. L. & Williams, J. M. (1986). Differential effects of sex and status on evaluation of coaching ability. *Research Quarterly for Exercise and Sport*, 57, 53-59.
- Perry, D. G. & Bussey, K. (1979). The social learning theory of sex difference: Imitation is alive and well. *Journal of Personality and Social Psychology*, 37, 1699-1712.
- Piaget, J. (1929). *The child's conception of the world*. London: Routledge.
- Plaisted, V. (1995). Gender and Sport. In T. Morris and J. Summers (Eds.), *Sport Psychology: Theory, Applications and Issues*. Milton, Qld, Australia: John Wiley & Sons.
- Rabbie, J. M., Schot, J. C. & Visser, L. (1989). Social identity theory: A conceptual and empirical critique from the perspective of a behavioural interaction model. *European Journal of Social Psychology*, 19, 171-202.
- Richardson, J. T. E. (1996). *Handbook of Qualitative Research Methods for Psychology and the Social Sciences*. Leicester, UK: British Psychological Society.
- Robson, C. (1993). *Real World Research*. Oxford: Blackwell Publishers.
- Rosenkrantz, P. S., Vogel, S. R., Bee, H., Broverman, I. K. & Broverman, D. M. (1968). Sex-role stereotypes and self concepts in college students. *Journal of Consulting and Clinical Psychology*, 32, 287-295.
- Ruble, D. N. & Ruble, T. L. (1982). Sex stereotypes. In A. G. Miller (Ed.), *In the Eye of the Beholder: Contemporary issues in stereotyping*. New York: Praeger.

- Russell, D. (1982). The Causal Dimension Scale: A measure of how individuals perceive causes. *Journal of Personality and Social Psychology*, 42, 1137-1145.
- Schoenfeld, W. N. (1942). An experimental study of some problems relating to stereotypes. *Archives of Psychology*, 270.
- Scruton, S. (1992). *Shaping Up to Womanhood: Gender and girl's physical education*. Buckingham, UK: Open University Press.
- Seavey, A. A., Katz, P. A. & Zalk, S. R. (1975). Baby X: The effect of gender labels on adult responses to infants. *Sex Roles*, 1, 103-109.
- Sherif, M., Harvey, O. J. , White, B. J., Hood, W. R. & Sherif, C. W. (1961). *Intergroup conflict and cooperation: The Robbers' Cave experiment*. Norman: University of Oklahoma Book Exchange.
- Sherriffs, A. C. & McKee, J. P. (1957). Qualitative aspects of beliefs about men and women. *Journal of Personality*, 25, 451-464.
- Signorella, M. L., Bigler, R. S. & Liben, L. (1993). Development differences in children's gender schemata about others: A meta-analysis review. *Developmental Review*, 13, 106-126.
- Smith, J. K. and Heshusius, L. (1986). Closing down the conversation: The end of the quantitative-qualitative debate among educational enquirers. *Educational Researcher*, 15, 4-12.
- Spence, J. T. & Helmreich, R. L. (1978). *Masculinity and femininity: Their psychological dimensions, correlates and antecedent*. Austin, TX: University of Texas Press.

- Sports Council. (1988). *Children's exercise, health and fitness: fact sheet*. London: Sports Council.
- Sports Council. (1992). *Sport in the nineties - new horizons: A draft for consultation*. London: Sports Council.
- Sports Council. (1993). *Young people and sport: Policy and frameworks for action*. London: Sports Council.
- Sports Council. (1994a). *The Brighton Declaration on Women and Sport*. London: Sports Council
- Sports Council. (1994b). *Women and sport: Policy and frameworks for action*. London: Sports Council.
- Sports Council & Health Education Authority. (1992). *Allied Dunbar national fitness survey*. London: Sports Council.
- Sports Council Northern Region. (1992). *Women in sport: Women in leadership in the northern region*. London: Sports Council.
- Sudman, S. & Bradburn, N. M. (1974). *Response Effects in Surveys*. Chicago: Aldine
- Swim, J. K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *Journal of Personality and Social Psychology*, 66, 21-36.
- Swim, J., Borgida, E., Maruyama, G. & Myers, D. G. (1989). Joan McKay versus John McKay: Do gender stereotypes bias evaluations?. *Psychological Bulletin*, 105(3), 409-429.

- Tajfel, H. (1969). The cognitive aspect of prejudice. *Journal of Social Issues*, 25, 79-97.
- Tajfel, H. (1981). Social stereotypes and social groups. In J. C. Turner & H. Giles (Eds.) *Intergroup Behaviour*. Oxford: Blackwell.
- Tajfel, H., Flament, C., Billig, M. G. & Bundy, R. F. (1971). Social categorisation and group behaviour. *European Journal of Social Psychology*, 1, 149-177.
- Tajfel, H. & Wilkes, A. L. (1963). Classification and quantitative judgment. *British Journal of Psychology*, 54, 101-114.
- Tester, N. (1998, June/July). Girl power. *Report: The magazine from the Association of Teachers and Lecturers*. 10-11.
- Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33, 529-554.
- Triandis, H. C. (1971). *Attitude and Attitude Change*. London: John Wiley and Sons.
- Turner, J. C. (1988). In M. A. Hogg & D. Abrams (Eds.), *Social Identification: A social psychology of intergroup relations and group processes*. London: Routledge.
- Turner, J. C. & Oakes, P. J. (1989). Self-categorisation theory and social influence. In P. B. Paulus (Ed.) *The psychology of group influence*. Hillsdale, NJ: Erlbaum.
- Urberg, K. A. (1982). The development of the concepts of masculinity and femininity in young children. *Sex Roles*, 8, 659-668.

- Wachs, T. D. & Gruen, G. E. (1982). *Early Experience and Human Development*. New York: Plenum.
- Wann, D. L. (1997). *Sport Psychology*. Upper Saddle River, NJ: Prentice Hall.
- Warshaw, P. R. & Davis, F. D. (1985). Disentangling behavioural intention and behavioural expectations. *Journal of Experimental Social Psychology*, 21, 213-228.
- Webb, E. J., Campbell, D. T., Schwartz, R. D. & Sechrist, R. D. (1969). *Unobtrusive Measures*. Chicago: Rand-McNally
- Webb, E. J., Campbell, D. T., Schwartz, R. D., Sechrist, R. D. & Grove, J. B. (1981). *Nonreactive Measures in the Social Sciences* (2nd ed.). Boston: Houghton Mifflin.
- Weinberg, R., Reveles, M. & Jackson, A. (1984). Attitudes of male and female athletes toward male and female coaches. *Journal of Sport Psychology*, 6, 448-453.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92, 548-573.
- Weiss, M. R. & Chaumeton, N. (1992). Motivational orientations in sport. In T. Horn (Ed.), *Advances in sport psychology*. Champaign, IL: Human Kinetics.
- Weiss, M. R. & Glenn, S. D. (1992). Psychological development and females' sport participation: An interactional perspective. *Quest*, 44, 138-157.
- Wicker, A. W. (1969). Attitudes vs. Actions: The relationship of verbal and overt behavioural responses to attitude objects. *Journal of Social Issues*, 25, 41-78.

- Wilder, D. A. (1978). Homogeneity of jurors: the majority's influence depends upon their perceived independence. *Law and Human Behaviour*, 2, 365-376.
- Williams, J. E. & Best, D. L. (1977). Sex stereotypes and trait favorability on the Adjective Check List. *Educational and Psychological Measurement*, 37, 101-110.
- Williams, J. M. & Parkhouse, B. L. (1988). Social learning theory as a foundation for examining sex bias in evaluation of coaches. *Journal of Sport and Exercise Psychology*, 10, 322-333.
- Wolfson, S., Ball, T. & James, I. (1985). Attitudes towards male and female sports participants. In J. Graydon (Ed.), *Proceedings of a one-day conference: Psychology, women and sport: Theoretical perspectives and practical applications*. Polytechnic of North London in conjunction with the Sports Psychology Section of the British Association of Sports Science.
- Wright, P. & MacLeod, H. (1992). [Edinburgh schoolchildren's sports participation]. Unpublished data.
- Yzerbyt, V. Y., Schadron, G., Leyens, J-P. & Rocher, S. (1994). Social judgeability: The impact of meta-informational cues on the use of stereotypes. *Journal of Personality and Social Psychology*, 66, 48-55.

APPENDIX A

Sample questionnaire from Study One

Steven Smith is a pupil at a school in Birmingham. He has one sister and one brother, who also go to the same school. His father is a fireman with the local fire brigade, and his mother is a nurse at a nearby hospital.

At school, Steven enjoys most subjects, but his favourite class is geography. One of Steven's proudest achievements was receiving a prize for the best Geography project in his school, for which he carried out a survey of his local area.

When asked about what kind of job he would like to do when he leaves school, Steven said "I think it is too early to make any decisions about jobs and things - I might go to college or something but I want to keep my options open for now."

Outside of school Steven enjoys going to the cinema, going ten-pin bowling, watching TV, and he spends at least two evenings a week reading. However his main hobby is playing the violin which he enjoys so much that he practises for an hour and a half every day, and regularly takes part in music competitions.

Go onto next page

Please answer the following questions - you may look back to the story if you wish.

1. How many people are there in Steven's family, including Steven?

2. What does Steven's father do?

3. a) What do you think Steven's main hobby is?

b) How much time does he spend doing it?

4. What subject does Steven like best at school?

Please wait - go onto next page when asked

Now, although you only know a little about Steven, I want you to imagine what kind of person you think he is. Read each of the words and phrases in this list and then indicate how well you think that word or phrase describes Steven.

Circle the **number 1** if you think the word is a **very poor** description of Steven.

Circle the **number 2** if you think the word is a **poor** description of Steven.

Circle the **number 3** if you think the word is **neither a good nor poor** description of Steven.

Circle the **number 4** if you think the word is a **good** description of Steven.

Circle the **number 5** if you think the word is a **very good** description of Steven.

If there is a word or phrase which you do not understand write a cross next to it, and move onto the next one.

	Very poor description			Very good description	
Caring	1	2	3	4	5
Works hard	1	2	3	4	5
Competitive	1	2	3	4	5
Loves children	1	2	3	4	5
Confident	1	2	3	4	5
Makes decisions easily	1	2	3	4	5
Aggressive	1	2	3	4	5
Popular	1	2	3	4	5
Easily fooled	1	2	3	4	5
Shy	1	2	3	4	5
Reliable	1	2	3	4	5
Athletic	1	2	3	4	5

Go onto next page

	Very poor description			Very good description	
Would be a good leader	1	2	3	4	5
Sensitive to needs of others	1	2	3	4	5
Independent	1	2	3	4	5
Gentle	1	2	3	4	5
Lazy	1	2	3	4	5
Copes well on own	1	2	3	4	5
Loyal	1	2	3	4	5
Willing to take risks	1	2	3	4	5
Cheerful	1	2	3	4	5
Honest	1	2	3	4	5
Jealous of other people	1	2	3	4	5

Please wait - go onto next page when asked

In the story Steven won a prize for his geography project.
Can you think of some reasons why he won the prize?

Below are some possible explanations as to why Steven won the prize. You may think that some are more likely explanations than others. Try and imagine what kind of person Steven is, and then indicate how likely or unlikely **you** think that explanation is to be true. Circle only one number for each explanation.

Circle the **number 1** if you think it is a **very unlikely** explanation.

Circle the **number 2** if you think it is an **unlikely** explanation.

Circle the **number 3** if you think it is **neither a likely nor unlikely** explanation.

Circle the **number 4** if you think it is a **likely** explanation.

Circle the **number 5** if you think it is a **very likely** explanation.

	Very unlikely explanation			Very likely explanation	
Steven tried very hard	1	2	3	4	5
Steven was lucky	1	2	3	4	5
Geography is not a difficult subject	1	2	3	4	5
Steven is good at geography	1	2	3	4	5
Steven is a hard worker	1	2	3	4	5
Steven was very fortunate on this project	1	2	3	4	5
Steven's parents encouraged him to work hard on this project	1	2	3	4	5
Steven finds geography an easy subject	1	2	3	4	5
Steven has been brought up to put maximum effort into his schoolwork	1	2	3	4	5
Steven is just a lucky kind of person	1	2	3	4	5
Steven has always had a natural talent for geography	1	2	3	4	5

Go onto next page now

Now please fill in the answers to the following questions about yourself.

Age _____

Sex (male or female) _____

Think about the hobbies, pastimes and sports that you do.

Now list below the ones you spend most time on, and tick the box to indicate how often you do them on average: almost every day, about 2 or 3 times a week, once a week, or once every 2 weeks.

Start with the activity which you do most often.

	almost every day	2 or 3 times/ week	once a week	once every 2 weeks
1. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Thank you very much for filling this questionnaire in.

If there any comments you would like to make about the story or the questions afterwards write them here.....

APPENDIX B

Sample questionnaire from Study Two

Rebecca Smith, aged 15, a pupil at West Holden Comprehensive School, has just been selected to take part in a summer school organised by Cambridgeshire Schools Rugby Union Association. Along with 45 other girls from around the county, Rebecca will spend three weeks at Foresham Hall. Each day the rugby players will practise the skills and techniques essential to rugby, and follow a challenging fitness training programme. Rebecca said "I am really looking forward to it. Three weeks of hard training will be tough, but I am sure that I will learn a lot from it". Rebecca was unsure how far she would go in the sport. "Well it's a bit early to say. I hope to play for the county team one day. Some of the girls at the summer school will be very good, and I'll need to work hard to keep up with them". The coaches who organised the summer school see it as an important way of improving standards throughout the sport. "It's great to get them all together," said one coach, "that way they all learn from each other, and because they are so competitive they push each other to do their best. It can only be good for the sport of rugby."

Now, although you only know a little about Rebecca, I want you to imagine what kind of person you think she is. Read each of the words below and then indicate how well you think that word describes Rebecca.

Circle the **number 1** if you think the word is a **very poor** description of Rebecca

Circle the **number 2** if you think the word is a **poor description** of Rebecca

Circle the **number 3** if you think the word is **neither a good nor poor** description of Rebecca

Circle the **number 4** if you think the word is a **good description** of Rebecca

Circle the **number 5** if you think the word is a **very good** description of Rebecca

If there is a word which you do not understand write a cross next to it and move onto the next one.

	very poor description			very good description	
Independent	1	2	3	4	5
Confident	1	2	3	4	5
Competitive	1	2	3	4	5
Strange	1	2	3	4	5
Attractive	1	2	3	4	5
Masculine	1	2	3	4	5
Gentle	1	2	3	4	5
Honest	1	2	3	4	5
Friendly	1	2	3	4	5
Selfish	1	2	3	4	5
Feminine	1	2	3	4	5
Bigheaded	1	2	3	4	5
Hard working	1	2	3	4	5
Obsessed	1	2	3	4	5

	very poor description			very good description	
Aggressive	1	2	3	4	5
Daring	1	2	3	4	5
Unusual	1	2	3	4	5
Athletic	1	2	3	4	5
Trendy	1	2	3	4	5

Section 2

Why do you think Rebecca was selected for the summer school? Here are some possible explanation why Rebecca was selected for the summer school. You may think that some of them are more likely than others. Try and imagine what kind of person Rebecca is and then indicate how likely or unlikely you think each explanation is to be true. Circle only one number for each explanation.

Circle the **number 1** if you think it is a **very unlikely** explanation.

Circle the **number 2** if you think it is an **unlikely** explanation

Circle the **number 3** if you think it is **neither a likely nor unlikely** explanation

Circle the **number 4** if you think it is a **likely** explanation

Circle the **number 5** if you think it is a **very likely** explanation

	Very unlikely explanation			Very likely explanation	
She was the best rugby player in her school.	1	2	3	4	5
Her parents wanted her to go.	1	2	3	4	5
Her parents are good friends with the summer school coaches.	1	2	3	4	5
There was a place for anybody who wanted to go	1	2	3	4	5

	Very unlikely explanation				Very likely explanation
No other girl in her school wanted to go.	1	2	3	4	5
She was lucky	1	2	3	4	5
She is a natural all-round sports player	1	2	3	4	5
She tries hard at all sports	1	2	3	4	5
She was in the right place at the right time	1	2	3	4	5

Section 3

Why do you think Rebecca wanted to go to the summer school? Here are some possible explanations as to why Rebecca wanted to go to the summer school. Again, try and imagine what kind of person Rebecca is, and then indicate how likely or unlikely each explanation is.

	Very unlikely explanation				Very likely explanation
She enjoys physical exercise.	1	2	3	4	5
She wants to become a better rugby player.	1	2	3	4	5
It's one step towards her ambition of playing for the county team.	1	2	3	4	5
Her parents encouraged her to go.	1	2	3	4	5
Her older brother has been to a summer school before.	1	2	3	4	5
She wanted a holiday in the country.	1	2	3	4	5

	Very unlikely explanation					Very likely explanation				
Her father was a rugby player and she wants to follow in his footsteps.	1	2	3	4	5					
She thinks that being sporty will make her popular at school	1	2	3	4	5					
She has some friends who are also going to the summer school	1	2	3	4	5					
She hopes to make new friends who also play rugby	1	2	3	4	5					

Section 4

In this section I want you to try and imagine what kind of person Rebecca is and then answer the following questions...

a) How much do you like Rebecca?

not at all					a lot
1	2	3	4	5	

b) Would you like her as a friend?

not at all				very much
1	2	3	4	5

c) How hard do you think she works at her schoolwork?

doesn't work at all				works very hard
1	2	3	4	5

d) How well do you think she will do in her exams?

will fail everything				will pass everything
1	2	3	4	5

e) Do you think she would be a good captain for her team?

very bad captain					very good captain
1	2	3	4	5	

f) Do you think Rebecca should do more work at rugby or less work at rugby?

a lot less work					a lot more work
1	2	3	4	5	

g) How many hours a week do you think Rebecca should spend training?

.....hrs

h)Do you think Rebecca should spend more or less time.....

	A lot less time				A lot more time
	1	2	3	4	5
Training?					
Doing homework?					
Going out with friends?					
Watching TV?					

I) How likely do you think it is that one day Rebecca will.....

	very unlikely				very likely
	1	2	3	4	5
play for the county team					
get married					
go to university or college					
have children					
become a well known sports star					

	very unlikely			very likely	
give up her sport when she leaves school	1	2	3	4	5
represent her country at rugby	1	2	3	4	5
become a rugby coach	1	2	3	4	5

Section 5

Can you think of some reasons why Rebecca plays rugby? Here are some possible explanations why Rebecca may play rugby. You may think that some explanations are more likely than others. Try and imagine what kind of person Rebecca is and then indicate how likely or unlikely you think each explanation is to be true.

	Very unlikely explanation			Very likely explanation	
She does it to keep fit	1	2	3	4	5
She enjoys competitions	1	2	3	4	5
She enjoys physical exercise	1	2	3	4	5
The training improves her good looks	1	2	3	4	5
Being sporty means she is popular at school	1	2	3	4	5
A lot of her friends also play rugby so she wants to join in too	1	2	3	4	5

Section 6

Now, fill in the following details about yourself...

Your age (in years).....

Your sex (male/female).....

Your class.....

How active in sport would you say you are?

Not at all
active in sport

Very active
in sport

1 2 3 4 5

If you do any sports outside of P.E. classes can you list them here and indicate how often you do them. If you do more than 3 sports just list the ones you do most often. If you do a sport which has a particular season (e.g. summer sports) then indicate how often you play that sport during the season.

	At least once a week	At least once a fortnight	At least once a month
1. _____	1	2	3
2. _____	1	2	3
3. _____	1	2	3

For the sport which you do most often indicate here the level at which you participate (e.g. play with friends, play for school team, play for county team, etc).

Level.....

Thank you for you help in completing this questionnaire. If you have any comments about the story or the questions afterwards then please write them here.

APPENDIX C

Sample questionnaire from Study Five

Please read the following story. Afterwards you will be asked to describe the main character in the story.

There is not much information in this story so read it slowly and carefully.

Daniel Smith, aged 15, a pupil at West Holden Comprehensive School, has just been selected to take part in a summer school organised by Cambridgeshire Schools Netball Association. Along with 45 other boys from around the county, Daniel will spend three weeks at Foresham Hall. Each day the netball players will practise the skills and techniques essential to netball, and follow a challenging fitness training programme. Daniel said "I am really looking forward to it. Three weeks of hard training will be tough, but I am sure that I will learn a lot from it." Daniel was unsure how far he would go in the sport. "Well it's a bit early to say. I would like to play for the county team one day. Some of the boys at the summer school will be very good, and I will need to work hard to keep up with them."

One of the coaches organising the summer school said "Daniel is very keen and eager to learn, and a great example to the rest of the boys. If you want to succeed at this sport you have to be prepared to work hard. Of course we want everybody to have fun too, and I am sure that Daniel will enjoy his three weeks with us."

One of the PE teachers at Daniel's school said "I am really pleased for Daniel. He has worked really hard all year and going to the summer school is a great reward. He is a fine example to the rest of the school." Daniel's geography teacher joked "I wish he worked as hard at his schoolwork as he does at netball!"

There was not much information in this story so you may like to read it over again.

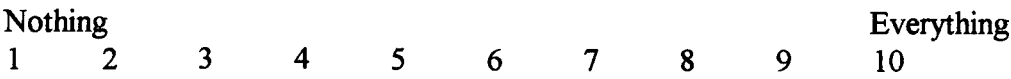
When you are ready turn over

There was not much information about Daniel in this story.

Can you indicate on this scale **how much you know about Daniel?**

The scale ranges from 1, which means that you know nothing about him at all, to 10 which means that you know everything about him.

Remember to circle only one number.



When you are ready turn over.

Instructions

READ THE FOLLOWING INSTRUCTIONS CAREFULLY

In this section describe what Daniel is like by indicating how well you think each word describes him.

Example:

The word is 'Happy'. The story may not tell you whether Daniel is happy or not, but you can guess whether he is or not from the picture you have built up of him. You might think that Daniel is happy and that this word is a very good description of him. In this case you would circle the number 5 like this....

Happy 1 2 3 4 5

If you think the word is a **very good description** of Daniel circle the **number 5**

If you think the word is a **good description** of Daniel circle the **number 4**

If you think the word is **neither a good nor a poor description** of Daniel circle the **number 3**

If you think the word is a **poor description** of Daniel circle the **number 2**

If you think the word is a **very poor description** of Daniel circle the **number 1**

REMEMBER FOR EACH WORD CIRCLE ONLY ONE NUMBER

If you want you can read the story and the instructions again.

When you are ready turn over and answer the questions.

Section 1

	Very poor description of Daniel				Very good description of Daniel
Competitive	1	2	3	4	5
Strange	1	2	3	4	5
Tough	1	2	3	4	5
Attractive	1	2	3	4	5
Friendly	1	2	3	4	5
Hard working	1	2	3	4	5
Masculine	1	2	3	4	5
Unusual	1	2	3	4	5
Sissy	1	2	3	4	5
Trendy	1	2	3	4	5
Aggressive	1	2	3	4	5
Athletic	1	2	3	4	5
Feminine	1	2	3	4	5
Gentle	1	2	3	4	5
Rough	1	2	3	4	5
Butch	1	2	3	4	5

Check that you have only circled one number for each word.

When you are ready turn over

Section 2

In this section answer each question by circling one number.

a) How much do you like Daniel?

not at all				a lot
1	2	3	4	5

b) Would you like him as a friend?

not at all				very much
1	2	3	4	5

c) How well do you think he will do in his exams?

will fail everything				will pass everything
1	2	3	4	5

d) Do you think Daniel should spend more time or less time on his sport?

a lot less time				a lot more time
1	2	3	4	5

e) How likely do you think it is that one day Daniel will...

	very unlikely			very likely	
play for the county team	1	2	3	4	5
get married	1	2	3	4	5
go to college	1	2	3	4	5
have children	1	2	3	4	5
become a famous sports star	1	2	3	4	5
become a coach	1	2	3	4	5
represent his country at netball	1	2	3	4	5
give up netball	1	2	3	4	5

Section 3

Why do you think Daniel plays netball? Below are some possible explanations - for each one indicate how likely or unlikely you think it is.

	very unlikely explanation			very likely explanation	
He does it to keep fit	1	2	3	4	5
He enjoys competitions	1	2	3	4	5
He enjoys physical exercise	1	2	3	4	5
He loves all sports	1	2	3	4	5
He enjoys being different	1	2	3	4	5
The training improves his good looks	1	2	3	4	5
He wasn't good enough to get into any other sports team	1	2	3	4	5
Being in the netball team means he is popular at school	1	2	3	4	5
He loves netball	1	2	3	4	5
A lot of his friends play netball so he wants to join in	1	2	3	4	5

Make sure that you have only circled one number on each line.

When you are ready turn over to the next page.

Section 4

What do you think Daniel's friends and classmates think about him playing netball?
Below are some possible ways in which his friends and other people might react. For each one indicate whether you think it is likely or not.

	very unlikely reaction				very likely reaction
His friends are jealous of his success	1	2	3	4	5
His parents are proud of him	1	2	3	4	5
The boys in the class call him names	1	2	3	4	5
His parents wish he would do more schoolwork and less netball	1	2	3	4	5
Some people laugh at Daniel behind his back	1	2	3	4	5
His friends encourage him	1	2	3	4	5
The girls in his class call him names	1	2	3	4	5
His friends think netball is boring	1	2	3	4	5
His friends want him to try a different sport	1	2	3	4	5

Make sure that you have only circled one number on each line.

When you are ready turn over to the next page.

Now that you have answered the questions can you indicate again **how much you know about Daniel?**

The scale ranges from 1, which means that you know nothing about him at all, to 10 which means that you know everything about him.

Remember to circle only one number.

Nothing										Everything
1	2	3	4	5	6	7	8	9	10	

Thank you for completing this questionnaire.

Now fill in the following details about yourself...

Your age (in years).....

Your sex (male or female).....

If you have any comments about the story or the questions write them here....