INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

The Effects of the Gymnast's Body Shape on the Judging of Gymnastics

by
Sheri Valiquette

A Thesis Submitted to the Faculty of Graduate Studies and Research in
Partial Fulfilment of the Requirements of the Degree of
Master of Arts (Education)

Department of Physical Education

Division of Graduate Studies and Research
Faculty of Education
McGill University
Montreal, Québec, Canada

July, 1996

© Sheri Valiquette, 1996.

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre reference

Our file Notre reference

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-29575-3

Canada

Abstract

The purpose of this study was to discover whether the gymnasts' body shape affects the judging of gymnastics. The goal was to discover whether the judges' perceptions of the gymnasts' body shape related to the gymnasts' performance scores. A second goal was to discover whether performance scores differed when given by judges with two levels of experience (experienced and novice). Videotapes of compulsory routines performed by 16 gymnasts on bars, beam and floor were constructed to simulate a competition. Eighteen judges, nine novice and nine experienced, evaluated the routines. They were also required to rate their perception of the aesthetic appeal of the 16 gymnasts' body shape. Each gymnasts' body mass index (BMI) was calculated from their height and weight measurements to quantify body shape. The scores were analyzed using repeated measures analysis of variance. The results showed a significant main effect of BMI on judges' perception of gymnasts' body shape scores, $F(3,48) = 9.50$, $p < .001$. Gymnasts with a medium-low BMI received significantly higher perception scores than gymnasts with a medium-high BMI and gymnasts with a high BMI. The results also showed a significant main effect ($p < .001$) of BMI on performance scores. Gymnasts with a high BMI received higher performance scores than those gymnasts with a low BMI, medium-low BMI and medium-high BMI. The effect of judge experience was very close to significant ($p < .059$) in relation to performance scores. Experienced judges gave lower performance scores than novice judges. Although all judges preferred smaller body shape, this preference did not affect the performance scores award by the judges.

Résumé

Cette recherche avait pour objectif de vérifier si le physique des gymnastes avait une influence sur l'évaluation de leur gymnastique. Le but était d'examiner si les perceptions du physique des gymnastes par les judges avaient un lien avec les points accordés aux routines des gymnastes. En second lieu, il fallait trouver les tendances sous-jouentes des points accordés aux routines des gymnastes avec différents IMC attribués par les judges selon deux niveaux l'expérience (novice et expert). Les enregistrements, sur vidéo, de routines obligatoires réalisées par les seize gymnastes sur les barres, la poutre et au sol, étaient réalisés de façon à immiter la competition. Dix-huit juges, neuf novices et neuf experts, ont évalué les routines. On leur a aussi demandé d'évaluer leur perception de l'attrait physique des 16 gymnastes. L'indice de masse corporelle de chaque gymnaste était calculé à partir des mesures de leur grandeur et de leur poids. Les résultats ont été analysés en mesurant de façon repétitive l'évaluation des désaccords. Les résultats ont surtout indiqué que les gymnastes avec un faible IMC ont obtenu une perception de points plus élevée que les gymnastes avec un IMC élevé. On a remarqué un lien significatif ($p < .001$) entre les points accordés aux routines et à l' IMC. Les gymnastes qui avaient un IMC élevé ont obtenu des points plus élevés pour leur routines quel les gymnastes qui ont un IMC bas. Le lien entre les points accordés aux routines et la competence des juges a été évalué assez révélateur ($p < 0.059$). Les judges experts ont donné des points plus bas aux routines que les judges débutants.

Acknowledgments

I would like to thank the Physical Education department for giving me the chance and experience to teach gymnastics and basketball to undergrad students who gave me variety and confidence in each day.

Special thanks is given to Diane Ste-Marie, who in my fourth year of undergrad gave me the experience and the confidence to pursue my Masters.

I would also like to thank all my excellent friends who made my time in Montreal enjoyable and memorable.

Love and appreciation to my brother and housemate, Dave who put up with my moods, played frisbee with me and was very consistent in helping review my thesis.

Special love and gratitude is given to my crutch, my support through everything, my mom. Without her encouragement and constant love, I would not have come this far.

A heartfelt thanks to my advisor, Dr. Downey, who with kindness and support guided me through with ease. Her patience and understanding was plentiful every step of the way.

Table of Contents

Page

ABSTRACT i

RESUME ii

ACKNOWLEDGMENTS iii

TABLE OF CONTENTS iv

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER I

INTRODUCTION 1

Subjectivity in Judging Sport Performance 1

Judging Gymnastics 2
Early studies 2
Judging biases 3
The aesthetic aspects of gymnastics judging 4

Health Concerns 6
Historical foundations 6
Implications 7

Judge's Experience 9

Nature and Scope of the Problem 11

Statement of the Problem 12

Operational Definitions 13

Delimitations 14

CHAPTER II

REVIEW OF LITERATURE 15

Changes in the Gymnasts' Physique 16

Research on Gymnast's Body Composition 18
Changes in the gymnasts' physique 18
Body composition of the gymnasts 19
Summary 22

Nutritional Concerns 22
Physiological concerns 22
Psychological concerns 24
Summary 28

Bias in Subjectively Judged Sports 29

Biases in Gymnastics Judging 30
International bias 30
Within-team order bias 31
Memory bias 33
Summary 35

Potential Body Shape Bias 36

Judges' Experience 37
Gymnastics studies 37
Summary 39

Summary of Chapter 40

CHAPTER III

METHODOLOGY 42

Participants 42
Gymnasts 42
Judges 43

Instrumentation 43
Videotaping 43
Selection of routines to be judged 44

Sequencing of routines....46
Gymnasts in static position....47

Procedures....48

Validity....49
Videotaped content being judged....49
Judging process....50

Pilot Study....50

Data Analysis....51

CHAPTER IV

RESULTS AND DISCUSSION....53

Judges' Perception of Gymnasts' Body Shape....53

Gymnasts' Performance Scores and Their Body Shape....57

Judges' Experience....61
Perception scores....61
Performance scores....63

Summary of Chapter....66

CHAPTER V

SUMMARY AND CONCLUSIONS....68

Methodology....68

Summary of Findings....69

Conclusions....70

Implications....71

Recommendations for Further Study....73

REFERENCES....75

APPENDICES....80

Appendix A: Camera Angles when Videotaping the Compulsory Routines....80

Appendix B: Development of An Instrument....82

Appendix C: The Compulsory Routine Judging Sheet....97

Appendix D: Assessing the Gymnasts' Body Shape Sheet....98

Appendix E: Verbal Instructions Given to the Judges at the Time of Testing....99

Appendix F: Judges' Experience Profile Sheet....100

Appendix G: Gymnasts' Age, Height, Weight and BMI....101

Appendix H: Distribution of Routines Across Apparatus and BMI Category....102

Appendix I: Consent and Information Form for Judges....103

Appendix J: Debriefing: A Bias for Body Shape....104

List of Tables

Table		Page
1	Changes in Age, Height and Weight of the U.S. Olympic Gymnast	17
2	Ranges of BMI Within Each BMI Category	45
3	Routines Performed by Gymnasts from Four BMI Categories	46
4	Distribution of Judges Across Sequential Orders of Routines	49
5	Mean Perception Scores for Body Shape of Gymnasts with Different BMI	54
6	ANOVA Summary for Judges' Perception Scores of Gymnasts' Body Shape for Aesthetic Appeal Scores as a Function of Judges' Experience and Gymnasts' BMI	55
7	Post-hoc Planned Comparisons of Mean Perception Scores of Gymnasts' Body Shape by BMI Category	55
8	Mean Performance Scores for Gymnasts with Different BMI	57
9	ANOVA Summary for Gymnasts' Performance Scores as a Function of the Order of Presentation of the Routines and the Judges' Experience	58

10 Post-hoc Planned Comparisons for Performance Scores by BMI Category.....59

11 Distribution of Gymnasts' Age Across the BMI Categories.....61

12 Judges' Years of Experience as a Judge, Athlete and Coach.....63

13 Mean Performance Scores Within Each Order for Experienced and Novice Judges.....66

List of Figures

Figure Page

1 Mean Perception Scores for Gymnasts' Body Shape for BMI Category 56

2 Mean Gymnasts' Performance Scores for BMI Category 59

3 Interaction of Judge Expertise and Order for Gymnasts' Performance Scores 64

CHAPTER 1

Introduction

The process of judging gymnastics has long been criticized for incidences of biases. Other sports such as figure skating, diving and boxing are also plagued with controversy over the lack of objectivity involved in the scoring process (Bruno, 1986; Seltzer, 1991; Whissell, Lyons, Wilkinson & Whissell, 1993). Memory bias (warm-up bias), within team order expectations, and international bias are a few prejudices which have been studied and documented for gymnastics (Ansorge & Scheer, 1988; Ansorge, Scheer, Laub & Howard, 1978; Ste-Marie, 1996; Ste-Marie & Lee, 1991; Ste-Marie & Valiquette, 1996). In addition to these, another bias may exist, originating from the gymnast's body shape (Falls & Humphrey, 1978).

Subjectivity in Judging Sport Performance

Many sports contain a certain amount of subjectivity in refereeing, umpiring and judging of the athletes. Others are regulated by more objective ways to determine the winner. Sports such as track and field have very objective ways of judging the winner, such as measuring throwing distances and determining winners by who crosses the finish line first. When the performance outcome is extremely close, computers and videotapes are used to make the accurate decision. Sports which require a medium amount of subjectivity, (e.g., hockey, football and basketball) rely on quasi-objective rules enforced by subjective referees. When dealing with quick decision calls, the referees or umpires may be swayed by certain influences such as the players'

suggestions, fans' exuberance and even the colours of the teams' jersey. There are also other sports plagued with subjectivity such as boxing, figure skating, diving and gymnastics, where the athlete's reward depends on the decisions from a panel of judges.

This study is concerned with the judging of gymnastics, which is primarily a subjective task of evaluating aesthetic movements. The gymnasts may be rewarded for their ability, or their scores may reflect other things (e.g. height or weight of the gymnmasts) that may be beyond the gymnast's control.

Judging Gymnastics

Typically, the duties of a judge of female gymnastics include watching between 30-50 routines in a one day competition on one of four apparatuses (vault, bars, floor or beam). While watching a gymnast perform, the judges must use shorthand to record the moves in the routine. At the same time, they deduct points for form and technique errors. When the routine is finished, they may also deduct points for rhythm and tempo, difficulty and other special requirements. When viewing compulsory routines where the skills and elements are the same for all competitors, the judges simply deduct points for form, technique, rhythm and tempo.

Early studies. Landers (1970) reviewed early studies that examined the reliability of gymnastics judging and attempted to objectify the scoring process. Averaging all the judges' scores (gross score) was compared to the conventional method of averaging the middle scores (net scores) to investigate changes in the individual standings. Results showed there was no difference in the individual

standings when using either method of scoring. Other studies discovered that events or meets which had small rather than large differences between judges' scores were more "objectively" judged than other events or meets.

Suggestions and proposals were made to attempt to control the subjectivity of judging gymnastics (Landers, 1970). Evaluation of filmed routines by the judges was suggested as a means of testing the judges' reliability and accuracy of judging. Reducing the judges' job to only one part of the exercise (difficulty, composition, or execution) was suggested as a way to increase the judge's accuracy and reliability. For higher performance levels, it was proposed that gymnastics adopt the system by which divers are judged. That is, multiply the difficulty score of the dive by the execution and composition total. In 1970, Landers suggested that "more energy needs to be expended in determining the amount of influence that the various social psychological factors such as the personal characteristics of the performer have on the accuracy of judges' ratings" (p.150). Since then these issues have been studied further.

Judging biases. Ste-Marie and Lee (1991) studied the effects of perceptual memory on gymnastics judging where prior exposure to routines might influence how the judge perceives a competitive performance. They examined two effects: (a) whether a single exposure to a move during warm-up would make an impression that would last until the actual competition, and (b) how this prior experience could bias the assessment of that move during the competitive performance. The emphasis of their experiment was on whether or not the judges were aware of the prior performance. A test simulating the warm-up/competition format was used, and it was

found that judges were influenced by having seen a prior performance of the skills. The judges were unable to control the influence of a prior exposure on their perception of the accuracy of a gymnastics move.

Ansorge and Scheer (1988) were interested in two kinds of bias patterns: (a) bias in favour of gymnasts from the judges' own countries, and (b) bias against gymnasts from foreign countries in close competition. The goal was to discover whether the judge scored the gymnast from his or her country higher, lower, or the same as the mean of the remaining judges. The results revealed the existence of the international bias in gymnastics judging with judges awarding higher scores to their own gymnasts.

Ansorge et. al. (1978) examined a bias that comes from the order in which gymnastics coaches place their gymnasts for competition. Because coaches typically place their gymnasts in rank order from poorest to best for competition, judges may expect that the quality of the routines will improve within a team order. The study did show this expectation bias. The judges expected that the routines at the beginning of the line-up would not be as good as the routines at the end of the order; so they were biased by this expectation when they evaluated the gymnasts.

The aesthetic aspect of gymnastics judging. The stereotype of the gymnasts' body has long been assumed and accepted. For example, consider two gymnasts: one is petit, 4'10" and weighs 90 pounds and the other is a little bigger, 5'1", and weighs 110 pounds. They can do the same skills, difficulty wise, and can execute them

exactly the same. Who would get the better score? A bias for body shape says that the smaller gymnast would because that gymnast is more pleasing to the eye.

Only one study has been found that considered the potential effect of the gymnast's aesthetic appearance on gymnastics judging. Falls and Humphrey (1978) looked at body type differences between placers and nonplacers in a regional collegiate gymnastics meet. Physical measurements were taken from 14 placers and 57 nonplacers in a meet of the Association for Intercollegiate Athletics for Women (AIAW) as well as from 54 reference non-athletic females. The findings suggested that certain body types may be a factor in championship performance scores in women's gymnastics. For example, Falls and Humphrey agree,

> The precise explanation for improved gymnastics performance among women with this body type is not available from the research conducted thus far. It is possible that the neat ectomesomorphic body type conveys a more favorable visual image to gymnastics judges and somehow influences the scores awarded (p.43).

It is conceivable that at elite levels the aesthetic component based on body shape affects the already subjective judging process (Warren, Stanton & Blessing, 1990; Falls & Humphrey, 1978). An example of this problem surfaced when the process for developing the U.S Olympic team was altered. Betty Okino, an athlete who was too injured to compete in either the U.S championships or the Olympic Trials was given another chance to become part the U.S Olympic Team. According to Ryan (1995), Betty's "far from flawless performance" (p.189) at this second trial won over Kim Kelly's perfect performance and forced Kim from the Olympic Team. Kim's coach recalls the selection committee's negative comments regarding Kim Kelly's

body type. The committee made comments about Kim's mature body of 5 feet 1 inch and 102 pounds. "We want to walk onto the floor in Barcelona petite and strong," (Ryan, p. 189) was one specific comment.

O'Brien (1991) gives an overview of bias in the judging of international elite gymnasts. He comments that there is a likelihood that the practice of bias is being taken back to the sport cultures of the countries present at championship meets and might be inserted at lower levels of gymnastics. This occurrence may be so for any bias in gymnastics.

Health Concerns

The possibility of a body shape bias in gymnastics raises concerns beyond those of equitable judging. Gymnastics is predominantly an anaerobic sport involving anaerobic training methods. It is not always easy to maintain a slender body while training for gymnastics. The gymnast may find it necessary to control her nutritional intake in order to maintain the petite physique that is favored at the present time and may be conducive to success (Calabrese, 1985; Ersoy & Gulgun, 1991; Reggiani, Arras, Trabacca, Senarega & Chiodini, 1989). Extreme behaviors, such as laxative over use and induced vomiting, may affect her health.

Historical foundations. Over the years the sport of gymnastics has changed drastically, thereby encouraging different body types. It is possible to trace the history of the gymnast's preferred body composition by recognizing some of the famous gymnasts (Bale & Goodway, 1990; Calabrese, 1985). In the mid 1970's Nadia Comaneci was slender and graceful, catering to the sport when rhythm, dance and

grace were the criteria. Mary Lou Retton had a stockier, more powerful and strong body to accomplish the powerful dynamic skills that were required to win in 1984. Most recent U.S. champions, such as Shannon Miller and Kim Zemeskial, possess petite, slender yet powerful bodies - a small body packed with power to produce the required highly technical twisting moves, and a slender look for the grace and rhythm that gymnastics now demands (Bale, 1994; Calabrese, 1985; Vercruyssen, 1984). The elite female gymnasts have a balanced mesoectomorphic physique with a slim body low in fat, small hips with relatively short trunks and quite broad shoulders (Beunen, Claessens, & Van Esser, 1981; Caladrone, Legise, Giametro, & Berlutti, 1986). Vercruyssen (1984) comments that there are biomechanical advantages of a short slim physique in rotational movements, arm support, and inverted balancing activities. The present sport of gymnastics has taken the past (Nadia Comaneci and Mary Lou Retton) to create gymnasts with a combination of power and grace such as Shannon Miller and Kim Zemeskial.

Implications. Incidences like Kim Kelly's displacement from the Olympic Team introduce the idea that a gymnasts' body shape may affect the score for her performance. Calabrese (1985) emphasizes the danger of this idea: "the aesthetic desirability for the prepubescent figures as well as the extreme youth of the competitors has introduced the pressures of serious athletic competition to an ever younger group of athletes, who are frequently unequipped to handle them" (p.24).

The nutritional intake that an athlete must concern herself with in order to maintain the "petite" physique that is conducive to success is an example of these

pressures (Calabrese, 1985; Ersoy & Gulgun, 1991; Reggiani et. al., 1989). A gymnast's body shape is mainly beyond her control since it is a genetic predisposition. However, many gymnasts attempt to control their body composition by controlling their food intake. They see this as a necessity in order to improve performance and to convey a more favorable image to the judges (Bale, 1994; Calabrese, 1985; Falls & Humphrey, 1978; Harris & Greco, 1990).

Harris and Greco (1990) emphasized that this type of pressure is particularly strong for athletes in sports such as gymnastics which require a thin and attractive appearance as part of the overall standard by which performances are judged (Borgen & Corgin, 1987; Calabrese, 1985; Rosen & Hough, 1988; Tveit, 1989). They asked a pool of gymnasts to explain more about this pressure. The open-ended questionnaire examined the reasons why the gymnasts were concerned about their weight and the everyday pressures to maintain body weight. Thirty-two percent of the gymnasts said the reason they feel the need to lose weight came from the need to look good in a leotard; 25%, from the belief that being thin helps one's performance; and 18%, from the emphasis on appearance. Fifty-six percent of the subjects reported they felt pressure from their coach to lose weight and 32% said the pressure came from others. Seven percent of the gymnasts reported that the pressure came from the fact that one is constantly watched and 18% said it came from the fact that they are constantly judged. "Several subjects commented that the pressure for gymnasts to be petite and to have a prepubescent figure makes it increasingly difficult to maintain an ideal shape for the sport as one matures" (Harris & Greco, p. 431).

Suddenly the gymnasts are engaged in another competition. This one is to stay slim and petite even though they may be doing harm to their bodies. There are no winners in this competition as young gymnasts acquire sometimes detrimental eating patterns that stay with them the rest of their lives (Harris & Greco, 1990; Rosen & Hough, 1988; Sundjot-Borgen, 1994; Ryan, 1995; Warren, Stanton, & Blessing, 1990). A number of studies examine weight concerns of athletes and specifically gymnasts (Sundgot-Borgen, 1994; Harris & Greco, 1990; Rosen & Hough, 1988; Warren et. al., 1990).

For the most part, gymnastics judges have a difficult job of subjectively scoring the gymnasts' routines. However, their perception of how a skill should look and how the gymnast should look doing the skill prevails. An ideal stereotype of the gymnasts' body has long been a prejudice that everyone has accepted. Maybe this expectation of the perfect "gymnast body" is too high a demand, especially if it is substituted for the actual talent of the gymnast.

Judge's Experience

Gymnastics judges receive extensive training throughout their careers. A beginner judging course is followed by theory and practical tests (Level 1). Then the judges must fulfill specific requirements each year to retain their certification (e.g., judge a number of meets at the provincial level). If they fulfill the requirements, the judges move up one level each year until reaching Level 3. Then they must take another course to advance further. The judge will gain more levels after the experience of another three years. After more training and testing, a judge reaches national and

international level as a Brevet judge. Over these years the judge is required to be actively judging and gaining experience. The judge develops a very keen eye for errors in technical performance. As the judges' level increases, so does the level of gymnastics they judge.

Expertise differences have been shown to affect the behaviour of athletes, teachers, coaches and judges (Armstrong & Hoffman, 1976; Bard, Fleury, Carriere, Halle, 1980; Franks, 1993; Imwold & Hoffman, 1983; Starkes & Deakin, 1984; Ste-Marie & Lee, 1991; Thomas, 1994). Armstrong and Hoffman (1979) investigated whether experienced tennis teachers could be differentiated from inexperienced tennis teachers on the basis of their ability to identify common performance errors of the tennis forehand . They found that the experienced teachers were more skilled at detecting errors than the inexperienced teachers.

Ste-Marie and Lee (1991) investigated whether specific prior processing effects would extend to the judgment of gymnastics moves and whether these effects were dependent upon the judge's expertise. It was expected that expert judges would be better at detecting errors than novice judges. They found that expert judges were, in fact, better than novices at identifying form errors.

Experienced gymnastics judges, having evaluated a higher level of gymnastics, may have developed expectations of a gymnast's body shape. It is possible that these expectations may affect the performance scores awarded. A difference in scores may exist between experienced and novice judges since novice judges may not have had the same experience in judging higher levels of gymnastics.

Nature and Scope of the Problem

Since judging is predominantly subjective in that one judge could be rewarding a gymnast for certain things while another is rewarding for others, it allows for the possibility of a bias for body shape. The book "Little Girls in Pretty Boxes, the making and breaking of elite gymnasts and figure skaters" (Ryan, 1995) has documented incidents of judges telling athletes that they need to lose weight in order to make it further in gymnastics. A bias for body shape says that the smaller gymnast would receive a better score because that gymnast is more pleasing to the eye.

Little research has been done on the existence of body shape biases at any level in competitive gymnastics. It is not known whether the petite gymnasts seem to score higher because they perform better or because of the judges' perception of an aesthetically pleasing gymnast. Only one study has ventured into the possibility that body shape may have influenced scores at a collegiate women's gymnastics championship (Falls & Humphrey, 1978). No research seems to have investigated whether this bias exists at the lower, club level of gymnastics. If this bias is prominent at the club level, then it may be important to make changes in the judging system or at least make judges, coaches and gymnasts aware of the potential problem. A better understanding of the issue could encourage judges to make a conscious effort to eliminate the bias. Coaches and gymnasts could begin to justify why a superior performance did not receive the score the routine deserved. As well, coaches could be conscious of the gymnasts' potentially dangerous attempts to become more aesthetically pleasing for the judges. Ultimately, recognition of this bias from the

countries' gymnastic federations could begin a reconstruction of the athlete's categories and other regulations and restrictions.

Statement of the Problem

The purpose of this study was to determine whether judges give higher scores to those gymnasts at the club level of competition who are shorter and weigh less. The study attempted to clarify whether the judges' perception of the aesthetic appeal of the gymnasts' body shapes relates to the scores that they give the gymnasts for performance of their routines. Finally, the study considered whether a bias for body shape differs for experienced judges (>10 years experience) in comparison with novice judges (<5 years experience).

The following hypotheses were examined:

(1) The judges' perception of the aesthetic appeal of a gymnast's body shape will be higher for gymnasts who are short and weigh less than for those who are taller and weigh more.

(2) The judges' scores for gymnasts' performances will be higher for gymnasts who are short and weigh less than for those who are taller and weigh more.

(3) Experienced judges will rate gymnasts lower for aesthetic appeal than novice judges.

(4) There will be a significant difference between performance scores awarded by experienced and novice judges.

Operational Definitions

Perception of aesthetic appeal: For this study aesthetic appeal indicates the judge's perception of how pleasing to the eye the gymnast's body shape is for competition at this lower club level of gymnastics.

Body shape: The gymnast's height and weight were measured and the shape of the gymnast's body was quantified by calculating her Body Mass Index using the following equation:

$$\text{Body Mass Index} = \frac{\text{Mass}}{(\text{Height})^2}$$

Present: Just before starting a routine and at the end of the routine, the gymnast facing the panel of judges stretches her body and arms to indicate the beginning and the end of her performance.

Compulsory routines: Routines which contain the exact same skills and dance moves and must be performed by every gymnast.

Form Errors: Body errors committed by the gymnasts in a routine. Examples of common form errors include unpointed toes and bent arms and legs when they are supposed to be straight.

Technical Errors: Errors in the actual whole movement of the skill. Examples of common technical errors include lack of amplitude or rotation of a skill.

Novice Judge: A certified judge with less than five years of judging gymnastics.

Experienced Judge: A certified judge that has had over 10 years of experience as a gymnastics judge.

Delimitations

1) Gymnasts that performed the compulsory routines were D level gymnasts from Quebec.

2) The judges were from the Eastern Region of Ontario.

3) The videotape of the compulsory routines only attempted to simulate the real competition setting for the judges.

CHAPTER 2

Review of Literature

Research on the sport of gymnastics and its athletes is restricted. Much of the literature originates from recent inquiries into concerns about the shift of elite gymnasts from women to young girls (Bale, 1994; Bale & Goodway, 1990; Bernadot & Czerwinski, 1991; Claessens & Veer, 1991; Novak, Woodward, Bestit & Mellerowicz 1977; Sinning & Wayne, 1978; Thorland & Johnson, 1981). Another portion of the research has investigated concerns about the eating patterns and the nutritional intake of the gymnasts. (Calabrese, 1985; Ersoy, 1991; Benson, Allemann, Theintz & Howald, 1990; Lindboe & Slettebo, 1984; Reggiani, et. al., 1989) Numerous studies have examined another recent interest regarding the increased tendency of gymnasts to engage in detrimental eating patterns (Harris & Greco, 1990; Petrie & Stoever, 1993; Rosen & Hough, 1988; Warren et. al., 1990; Sundgot-Borgen, 1994; Thorton, 1990).

There has been very little research on gymnastics itself and on the factors that affect the subjectivity of judging gymnastics. Potential biases in the judging process that have been researched thus far are political, within-team order and memory based (Ansorge & Sheer, 1988; Sheer & Ansorge, 1975; Ste-Marie, 1996; Ste-Marie & Lee, 1991; Ste-Marie & Valiquette, 1996). One study has ventured into the possibility of a judging bias due to the gymnast's body type (Falls & Humphrey, 1978). A review of this particular study and of studies from each of the categories mentioned above is given to clarify the nature of the problem and the motivation for this research.

Changes in the Gymnasts' Physique

The present study investigates the gymnasts' body shape and its possible influence on the judges' scores. It is necessary to examine the literature that explains the gymnast's past and present body compositions and shape. This literature is plentiful with studies on the body composition of both club and elite level gymnasts. The recent flurry of studies shows there are concerns about the health practices of gymnasts. This interest also reflects the need to understand the negative media that gymnastics has recently received.

In the mid seventies, the gymnastics world was romanced by Olympic medalist, Nadia Comeneci. She had a slender, graceful body that had a beautiful line when performing intricate skills. Gymnasts such as Kathy Johnson, Elfie Schlegel and Olga Korbut followed with bodies similar to Nadia's. In the late eighties and early nineties gymnasts smaller in stature, such as Kim Zemeskial, Shannon Miller and Dominique Dawes, took the floor. However, there were exceptions such as Mary Lou Retton who was very petite but stocky, showing us powerful skills and a dynamic spirit. More recently, the exception to this trend of smaller, younger gymnasts is Svetlana Boginskaya who, at the age of twenty-two, is still at the top of competition with a taller woman's body. Bela Karolyi says that the new American 1996 Olympic hopeful, 14 year old Dominique Moceanu, "has the looks, grace and style of Nadia Comaneci and the exuberance of Mary Lou Retton"(Guise, p.1996). This petite body shape is the more preferred body shape, even though the exceptions exist.

As a partial result of these changes in preferred body shape since Comaneci's victory, there has been a drastic decline in age and size among American gymnasts. Lincoln (1982) has pegged this as the "Nadia syndrome." In 1976, the U.S Olympic team's average age was 17.5 years old, and the average gymnast was 5 feet 3 inches and weighed 106 pounds. The U.S gymnastics team member for the 1992 Olympics averaged 16 years old, stood 4 feet 9 inches tall and weighed 83 pounds (see Table 1). That is, the average gymnast was one year younger, 6.5 inches shorter and 23 pounds lighter than those sixteen years before (Ryan, 1995). The elite female gymnasts now have a balanced mesoectomorphic physique with a slim body lower in fat, small hips with relatively short trunks and quite broad shoulders (Beunen et. al., 1981; Caladrone et. al., 1986).

Table 1

Changes in Age, Height and Weight of the U.S Olympic Gymnast

	1976	1992
Age	17.5	16
Height	5'3"	4'9"
Weight	106lbs	83lbs

The "code of points" has had great influence on these changes documented in the literature. The code of points is the gymnastics book of skills with their difficulty marks. Paralleling the continuous changes in the code of points, the present sport of

gymnastics has changed. As a result, the body shape of the past (Nadia Comeneci and Mary Lou Retton) has been conglomerated to create gymnasts with body shapes similar to Shannon Miller, Kim Zemeskial and Dominique Moceanu (Bernadot & Czerwinski, 1991). These gymnasts have petite, powerful, and flexible bodies that can perform the difficult skills and are very aesthetically pleasing to the eye.

Research on Gymnast's Body Composition

The gymnasts of today are shorter, lighter and have a lower percentage of body fat than gymnasts from the past and athletes from other disciplines. It has also been found that the age of gymnasts has decreased over the years.

Changes in the gymnasts' physique. Claessens, Veer, Stijnen, Lefevre, Maes, Steens, & Beunen (1991) took anthropometric measurements of outstanding male and female gymnasts from the 1987 World Championship gymnastics competition. Two hundred and one elite female gymnasts were measured for body dimensions and somatotype. The results were compared with reference data obtained at international gymnastics competitions of previous years. They found female gymnasts were significantly ($p < .01$) smaller than gymnasts of previous years in weight and in all length dimensions including height. A decline in mean height from 159cm to 154.3 cm was noticed and seemed to be accentuated between the 1976 Olympics and the World Championships at Budapest in 1983. Along with this height decline, there was a weight reduction from 52 kg to 45 kg. The skinfold measurements showed the 1987 gymnasts were thinner with 48% of the gymnasts showing an extremely low amount of

subcutaneous fat. Not only did these physiological factors differ, but the gymnasts in 1987 were younger than all previous gymnasts with a decline in mean age from 22.7 in 1964 to 16.5 years of age in 1987. The researchers also found that there was more change in the shape of the female gymnast from 1967 to 1987 than that of the male gymnast.

Body composition of the gymnasts. Thorland and Johnson (1981) examined body composition and somatotype characteristics of Junior Olympic athletes. Skinfolds were taken along with an underwater weighing for body composition. The athletes studied were gymnasts, wrestlers, divers, track and field jumpers, throwers, and middle distance runners. All categories of athletes were compared with each other and also with the reference non-athlete female. Gymnasts and divers were categorized together. The researchers found a significant difference between the somatotype of the gymnasts and divers compared to other athletes. Comparisons of the female athletes and the reference female revealed that the Junior Olympic competitors, except gymnasts and divers, were considerably taller. In addition, gymnasts and divers were lighter than throwers in body weight. All competitors except for gymnasts and divers were notably higher in lean body weight than the reference female. It appears that the gymnast is the exception and exists on the smaller line of the athletes (Thorland & Johnson, 1991).

The changes in the gymnasts' physique over the years could be linked to the change in age of the elite gymnast. Benardot and Czerwinski (1991) looked at selected body composition measures of female junior elite gymnasts. In their study they

compared 146 gymnasts that belonged to one of two age groups: a younger group (7-10 year olds) and an older group (11-14 year olds). By using age-related height:weight ratios, they found that gymnasts were in the 50th percentile for weight regardless of age. However, weight:age and height:age percentiles progressively dropped from the 48th to the 20th percentile as age increased. As the gymnasts grew older, they became progressively smaller in weight and height for their age but were highly muscled for their size.

Novak et. al. (1977) compared three groups of athletes: runners, swimmers and gymnasts. After taking physical measurements of age, height, weight and body fat percent, he found the runners to be significantly older than the swimmers and the gymnasts. All three groups were of similar height and weight. Even though the gymnasts seemed to be of smaller stature while carrying less weight, these differences were not statistically significant. Total body fat was found to be significantly higher in swimmers (18.9%) than in runners and gymnasts (12.9%).

Growth and sexual maturation of female gymnasts has also been examined. Claessens, Malina, Lefevre, Beunen, Stijnen, Maes, & Veer (1992) compared growth measurements and menarcheal status of elite female gymnasts with a reference sample of nonathletic girls. Two hundred and one gymnasts who participated in the 24th World Championship Artistic Gymnastics in 1987 were subjects for this study. Weight, stature, age at menarche, biacromial and bicristal breadths as well as the somatotype of gymnasts aged 13-20 were compared with the controls. Anthropometric dimensions of the gymnasts increased with age until about 16 years of age and then

tended to plateau. Compared with the reference nonathletes, the elite gymnasts are considerably shorter and lighter with narrower shoulders and hips. On average, elite gymnasts attained menarche at a much later age than nonathletes (gymnasts 15.6 ± 2.1 years and nonathletes 13.2 ± 1.2 years).

More specific to gymnastics and judging, Falls and Humphrey (1978) looked at body type differences between placers and nonplacers in a regional gymnastics meet. The physical measurements were done on 14 placers (P) and 57 nonplacers (NP) in an Association for Intercollegiate Athletics for Women (AIAW) meet as well as on 54 reference non-athletic females. The data were compared across the subjects and with data from previous studies. The 1978 gymnasts were taller and approximately 4.1 kg heavier than the Springfield College national championship gymnastics team previously studied in 1972 (Sinning & Lindberg, 1972). The gymnasts in the Falls & Humphey (1978) study were also taller and heavier than Olympic (de Garay et al., 1974; Hirata, 1966) and Russian (Carter, 1970) gymnasts previously studied. The position of the placers within the somatochart was virtually identical to that of the Olympic gymnasts. Although the differences were non-significant, both gymnast groups were shorter and lighter in weight than the non-athletes. All the differences of estimated body fat among these groups were significant ($F = 19.01$), and all but two of the P were below the mean of the NP for estimated fat percentage. The authors concluded that certain body types may be a factor concerning championship performance in women's gymnastics. "The precise explanation for improved gymnastics performance among women with this body type is not available from the

research conducted thus far. It is possible that the neat ectomesomorphic body type conveys a more favorable visual image to gymnastics judges and somehow influences the scores awarded" (Falls & Humphrey, 1978).

Summary. The face of gymnastics has changed. Until the 1970's the gymnastics world was dominated by women who emphasized the grace and artistry of gymnastic performance. Today female gymnasts are lighter, shorter and more ectomorphic than earlier gymnasts. Their physique is well suited to the biomechanical demands of a more acrobatic performance (Bale, 1994; Calabrese, 1985; Vercruyssen, 1984).

Nutritional Concerns

The statistics presented above are considered normal and would not be questioned if the physical measurements taken (height, weight, body fat percent and somatotype information) were genetic. However, not all athletes have the perfect gymnast body naturally. To what lengths do these gymnasts go to obtain or maintain the desired physique? The problem that exists is maintaining a low percentage of body fat while training in a predominantly nonendurance sport.

Physiological concerns. In 1985, Calabrese had 20 level A club gymnasts record the quantity and type of food, drink and nutritional supplements ingested during a three day training period. It was found that caloric consumption was marginal, at 99 per cent of the recommended daily allowance (RDA) defined by the Nutritional Data Base at Case Western Reserve University. These were normal values if compared with non-athletes who did not train intensively four to five times a week. In addition, the

nutritional density of the diets appeared low with subjects taking less than the RDA of vitamins A and D, folic acid, calcium, phosphorus, magnesium, and zinc. A similar study was done using child gymnasts (mean age 11.5 ± 0.5 yrs.) who had been involved in the sport for 3.9 (±.5) years (Ersoy, 1991). The gymnasts' body fat percentage was determined using the skinfold caliper in eight regions. The average height and weight of the subjects were 142.4 (±2.8) cm and 31.6 (±1.5) kg. The subjects' average fat percentage was 22 %, which was 6% higher than in other studies. The 20 gymnasts consumed only 59% calories of their RDA. Another discovery was that the gymnasts had an inadequate intake of calcium, iron, vitamin A, thiamin and niacin. In addition, 15 of the 20 gymnasts (75%), stated that they experienced episodes of dizziness, weakness and shortness of breath during their training periods (Ersoy, 1991).

Reggiani et al. (1989) stated that the leanness of gymnasts may be a result of genetics, since all their data seems to be within the normal standard for athletes their size. They studied 26 young female gymnasts with an average age of 12 years. Their average height, weight and body fat percentage were 145.8 (±8.5) cm, 37.9 (±6.9) kg and 15.0 (±3.5) % , respectively. Even though they found that the gymnasts' caloric intake was 25% less than what The National Institute of Nutrition recommends for same aged girls, it was still considered within the normal range. These gymnasts, however, had deficient dietary consumption of some nutrients. Particularly important was their lack of carbohydrate intake which measured 47.7% of the total daily intake vs. the recommended consumption of 55-60%.

These nutritional practices of low calorie diets without taking into consideration the expenditure of energy can cause serious health problems. These problems are often associated with female individuals engaged in sports such as ballet, figure skating, long distance running and gymnastics. Dieting, along with the inadequate intake of certain nutrients, can lead to delayed menarche, amenorrhoea, risk of osteoporosis, anorexia nervosa and other eating disorders (Bale, 1994; Claessens et. al., 1992).

Psychological concerns. It might be sufficient to say that proper monitoring of the gymnasts' nutritional intake should eliminate this problem. It should be the responsibility of the coaches to devise safe and proper diets that encompass all the gymnasts' nutritional needs, including the sufficient intake of calories for the amount of energy expenditure. The following studies examined the weight concerns of athletes and specifically gymnasts. Many of the studies showed that not only were coaches uninvolved in their athletes' diets, but they were pressuring the athletes in some way to engage in detrimental eating patterns.

Harris and Greco (1990) realized the pressure placed on female athletes to maintain low body weight and especially a low percentage of body fat. This pressure is particularly strong for athletes in sports such as gymnastics, which require a thin and attractive appearance as part of an overall standard by which performances are judged (Borgen & Corgin, 1987; Calabrese, 1985; Rosen & Hough, 1988; Tveit, 1989). Harris and Greco (1990) hypothesized that the unusual pressures on these gymnasts to perform and to maintain a thin but strong body would result in a higher reported use of

dangerous weight control behavior. Consequently, they thought that gymnasts would have higher scores on the Eating Disorder Index (EDI), the Drive for Thinness, Body Dissatisfaction, and Perfectionism scales, and on the Eating Attitudes Test (EAT).

Twenty-eight college level gymnasts were given these tests as well as an initial questionnaire asking a number of questions about their weight and how it was perceived by themselves and others. From the questionnaire (Falls & Humphrey, 1990), subjects reported using a variety of methods when trying to lose weight: 75% of those tested were on a mild diet, 71% increased their level of exercise to control weight, 43 % were on a strict diet, 18% were fasting and 7% had used laxatives. Fifty-six percent of the gymnasts reported they felt pressure from their coach to lose weight. They averaged 128 minutes a week thinking about their weight and 40 minutes talking about it. The open-ended section of the questionnaire examined the reasons why the gymnasts were concerned about their weight, and the everyday pressures to maintain body weight. The pressures came from several sources: 32% said the pressure was from others; 29%, from their coaches; 32%, from the need to look good in a leotard; 25%, from the belief that being thin helps one's performance; and 18% from the emphasis on performance. Seven percent said the pressure came from the fact that one is constantly watched, and 18% said it came from the fact that they are constantly judged. "Several subjects commented that the pressure for gymnasts to be petite and to have a prepubescent figure makes it increasingly difficult to maintain an ideal shape for the sport as one matures" (Harris & Greco, p. 431).

In an effort to identify risk of disordered eating patterns in competitive female athletes, Warren et. al., (1990) compared college aged female athletes to college female nonathletes by using the Eating Disorders Inventory (EDI) and the Eating Attitudes Test (EAT). Seventy-four athletes from gymnastics, cross-country running, basketball, golf, volleyball, swimming, and tennis took part. Fifty-two college females served as the non-athletes. An analysis was done on the two sports for which low body weight was deemed central to performance: cross-country running and gymnastics. It was found that gymnasts received higher scores on the EAT, and on the EDI subscales of Drive for Thinness, Body Dissatisfaction and the Introceptive Awareness. The rest of the analyses were performed on four groups: gymnasts (n=15), cross-country runners (n=12), athlete controls (n=47), and nonathlete controls (n=52). The results showed that groups differed on two EDI scales. On the Drive for Thinness, a Tukey test revealed that the gymnasts had higher scores than runners. Also, on the Body Dissatisfaction scale, gymnasts and nonathlete controls were more dissatisfied than runners. When the scores were compared with those studies specific to bulimics and anorexia patients, the researchers found that 20% of gymnasts, 17% of athlete controls, 12% of nonathlete controls and no runners received the set cutoff score for diagnosed anorexia and bulimic patients on the EAT test. Twenty percent of gymnasts, 4% of athlete controls, 10% of nonathlete controls, and no runners evidenced weight preoccupation. On the EDI personality scores, only the gymnasts' Perfectionism mean score of 7.27 approached those reported for bulimics in other studies (mean =7.8; Nudelman, Rosen, & Leitenberg, 1988). Warren et. al. (1990) questioned why

gymnasts might be more preoccupied with weight than runners, even though a lean physique is important for performance in both running and gymnastics. "A plausible hypothesis is that gymnasts may place more emphasis on an aesthetic standard for appearance. Judgment of gymnastics performance cannot be said to be as objective" (Warren, et. al. p.568).

Forty-two female gymnasts, aged 17-22 years, were given the Michigan State University weight control survey in Rosen and Hough's study (1988). This survey helps the clinician or researcher identify factors associated with the use of pathogenic weight-control methods. One or more of the following behaviors constitutes pathogenic weight-control behavior: self-induced vomiting, use of laxatives or diuretics for weight loss, regular use of diet pills, fasting for more than one day, and/or fluid restriction for at least one day per week (Rosen & Hough, 1988). If the gymnasts employed any of these techniques at least twice weekly over a period of three or more months, it was recorded that she was using a pathogenic weight control method. Fluid restriction was recorded if it had occurred at least once a week. All 42 of the gymnasts were actively dieting. Half the gymnasts said they dieted in order to improve their performance, the other half dieted in order to improve their appearance. Of the total group, 62% were using at least one form of pathogenic weight control. Self-induced vomiting, the use of diet pills, and fasting were the most frequently used methods. Twenty-eight of the 42 gymnasts were told by their coaches they were too heavy. Seventy-five percent (n=21) of these gymnasts proceeded to use pathogenic weight

control methods. However, of the gymnasts that were not told they were overweight, 36% (n=5) also used pathogenic weight-control techniques.

Use of pathogenic weight-control methods can adversely affect athletic performance and threaten health. Such methods can result in malnutrition, dehydration, loss of vital electrolytes, hypoglycemia, and excessive adrenergic stimulation. Eventually, these problems can significantly decrease performance capabilities, heighten the risk of injury and even lead to death (Rosen & Hough, 1988).

A few harsh words from someone important in an athlete's life can do a lifetime of damage. The coach is someone whom the athlete is always trying to please. When the coach's comments focus on highly personal and sensitive issues, they can be devastating to the athlete's self-esteem. Proof of the results of these drastic comments lies in the personal interviews and stories that are told in "Little Girls in Pretty Boxes" (Ryan, 1995). Chapter two, "They stole her soul and they still have it," is about the eating disorders of famous gymnasts who, because of the nature of the sport, (among other things) lost not only their childhood but also their lives (p.95).

Summary. This review of literature emphasizes that many elite gymnasts are not consuming the proper amount of nutrients, especially for the amount of energy expenditure. Even more, some gymnasts are engaging in pathenogenic eating patterns to control their body shape. Research attempts to explain why these problems exist by examining the internal and external pressures placed on the gymnasts. The gymnastics federations in both the United States and Canada have confronted many of these concerns. Coaches have been educated, and a task force has been implemented for

parents and gymnasts so they might understand the dangers of dieting and the signs of eating disorders (Ryan, 1996). These are all direct ways of handling the problem.

Bias in Subjectively Judged Sports

The literature about observer bias indicates that people performing tasks having some degree of subjectivity are susceptible to biasing effects. In sports that require a panel of judges to evaluate the athletes' performance, previously formed expectation of performance outcome can determine the judgmental response (Brawley, Landers, Miller, & Kearns, 1979). Johnson (1953) explains "that one tends to see situations in ways that fit one's assumptions even to the extent of distorting or omitting detail" (Brawley et. al., p.15). Studies have been done on figure skating, diving and gymnastics to investigate and better understand the potential biases involved in these subjectively judged sports.

In 1993, Whissel et. al. looked at international bias in judgments from the 1984 and 1988 Olympic figure skating performances. They performed tests of probability on judges' scores and found that the judges rated skaters from their own country above the average scores of the remaining judges. The judges awarded skaters from their own country maximum scores and overall ranks that also tended to be higher than the skaters' final Olympic standings. It was concluded that an international bias exists at this level of figure skating.

In 1986, Bruno researched the possible scoring bias in springboard and platform diving. Specifically, he wished to examine the relationship between the degree of difficulty of a dive (DD) and the scores awarded for a dive. Using a

regression analysis, it was found that as divers attempted more difficult dives (increased DD), the judges were less inclined to award lower scores. Bruno (1986) went on to suggest the use of a Bayesian approach to revise the DD tables in order to reduce this scoring bias.

Biases in Gymnastics Judging

In the sport of gymnastics there are also numerous biases that can affect the subjectivity of scoring the gymnasts' routines. Those that have been researched are political bias, within team order and memory bias (Ansorge & Sheer, 1988; Sheer & Ansorge, 1975; Ste-Marie, 1996; Ste-Marie & Lee, 1991; Ste-Marie & Valiquette, 1996). It is documented that most biases are beyond the judge's control.

International bias. Ansorge and Scheer (1988) were interested in the bias patterns of men and women judges at the 1984 Olympic Games. Two kinds of bias patterns were of interest: 1) bias in favor of the gymnasts from the judges' own countries, and 2) bias against gymnasts from countries in close competition. Judging assignments and results were obtained from the 1984 Olympic gymnastics competition. The scores of each individual judge were compared with the mean of the remaining three judges on the panel. The goal was to discover whether the judge scored the gymnast from his or her country higher, lower, or the same as the mean of the remaining judges. The analyses showed that judges' scores were higher in 122 instances and lower 12 times than the score of other judges on the panel. The analysis also showed that the scores assigned by individual judges to the gymnasts in close competition were lower than the average of the remaining judges on the panels. This

study reveals evidence that supports the existence of international bias in gymnastics judging.

Ste-Marie (1996) researched this further to test whether international bias in gymnastic judging could be partially explained by unconscious influences in the form of exposure or perceptual fluency effects. It is generally assumed that international biases are a result of judges consciously awarding athletes from their politically affiliated countries higher scores than athletes from competing countries. However, Ste-Marie hypothesized that there is a possibility that unconscious influences contribute to this international bias and it may not be controlled by the judges' intentions. To test this, videotaped gymnastic routines were arranged so that they were seen once, twice or three times and viewed by gymnastics judges. It was expected that those routines seen more would be scored higher and liked better. This simulated the scenario that judges have had numerous occasions to evaluate routines by gymnasts from their countries before encountering the routine at an international competition. The results indicated that exposure repetitions of at least three presentations did not result in higher ratings of liking for a routine, nor higher scores. Therefore, no support was shown for the possibility of this potentially unconscious influence contributing to international bias.

Within-team order bias. Ansorge et. al. (1978) examined a possible bias stemming from the order in which gymnastics coaches place their gymnasts for competition. Coaches typically place their gymnasts in rank order from poorest to best for competition. Judges may expect the quality of the routines will improve within a

team order. To test whether gymnastics judges are influenced by within-team order, original videotapes of 111 gymnasts from five different teams were edited and the within-team orders were changed. On one tape a gymnast's routine would appear in the first position, and on the other tape the same gymnast's routine would appear in the fifth position. Each of the 10 judges viewed both tapes in two sessions separated by 48 hours. A significant difference between the scores appeared when the routines were presented in a different order. Two evaluations of the same routine resulted in higher scores for the routines when they were placed in the fifth position. This was an expectation bias. The judges expected that the routines at the beginning would not be as good as those at the end of the order. For this bias, there was a possible solution. Sheer and Ansorge recommended that the coach randomly distribute the gymnasts in order to eliminate this bias.

The authors of the previous study found that not all judges were influenced by this naturally induced expectation of the gymnast's performance. To investigate this performance expectation further, Sheer and Ansorge (1979) included internal and external locus of control of the judges. They wished to establish a relationship between the degree to which female judges were influenced by the within team position of gymnasts and their internal - external locus of control. People with an internal locus of control believe that their behaviors influence the reinforcements they receive. A person with an external locus of control tends to believe that their reinforcements are controlled by fate, chance or the power of other people. Results of this study indicated that judges with an external locus of control scored the gymnasts

in the fifth position higher than the same gymnasts when they were the first competitor. The judges with internal locus of control were not affected by the gymnast's position, even though they may have expected that the scores would improve within the team orders.

Memory bias. A memory bias which results from prior exposure to skills that are to be evaluated has been an issue of recent study. The factors involved in gymnastics judging and the effect of perceptual memory may impact on how the judges will later perceive events.

Ste-Marie and Lee (1991) found this true when they examined (1) the effects of prior exposures to a gymnastic move on the later assessment of performance and (2) whether or not the judges were aware of the prior performances. Eighteen judges viewed 80 gymnastics skills in the first phase and 64 skills in the second phase of the study. Two video clips of different gymnastics moves were made. One clip was the gymnast performing the skill perfectly, the other was the same gymnast performing the skill with a form error. An example of a form error would be the gymnast bending her legs when they should be straight. The video clips were edited so that in the first phase of the test, the judge might see a perfect performance of the move. The judges had to verbally express whether the skill was a perfect performance or whether a form error was committed. During the second phase, the judges might see the move performed exactly the same way or with a form error, or they might see a new move entirely. It was found that perceptual accuracy differed as a function of the relationship between an item in the first phase and its parallel item in the second test phase. More

specifically, moves with the same performance in both phases were perceived with the highest accuracy, new moves were perceived less accurately, and the lowest level of perceptual accuracy occurred for moves where the performance had differed between the two phases.

To put this into perspective for a gymnastics competition, the judges usually watch the gymnasts warm-up before they compete in the event, and this prior presentation can affect how the judge evaluates the actual competitive performance. If the gymnast performed perfectly in the warm-up but not so well in the actual competition, the judges' prior viewing of the perfect warm-up might positively influence their judging of the competition routine.

The experiment's (Ste-Marie & Lee, 1991) second emphasis was on whether or not the judges were aware of the prior performance. Gymnastics skills were presented in a third phase to test direct memory. The judges had to verbalize whether they have seen the skill in the previous phases by making an old new decision on 40 more skills. The judges were unable to control the influence of a prior exposure on their perception of a gymnastics move. That is, the judges were unaware of the indirect influence of the prior exposure on their perception of a gymnastics move.

It is also evident that perceptual experiences show long lasting effects. Ste-Marie and Valiquette (1996) conducted a similar test to that of Ste-Marie and Lee (1991). In this investigation the effects from prior experience were examined over three different delays (No delay, 24 hour delay, one week delay). The no delay condition simulated the competitive situation where the judge experiences the

gymnast's brief warm-up just before competition. Similarly, a 24 hour delay occurs in the situation of a two day competition, which includes an individual competition the first day and a final competition the second day. Finally, during elite competitions, gymnasts may need to perform routines throughout an entire week. In all of these cases, the judge's prior viewing may have affected her perceptions at a later time. Moves with the same performance were judged with the most accuracy, followed by the new moves. The different moves were judged with the least accuracy. This pattern of results was the same across all delay conditions with no decrease in magnitude. While recognition results (the aware memory) decreased over time, the perceptual result (unaware memory) stayed the same across time, showing that the dangerous bias found by Ste-Marie and Lee (1991) is evident even as time passes. As it stands now, the only way for the gymnast to benefit from this memory bias is to perform outstandingly in the first competition and in every warm-up. Restricting judges from watching the warm-ups and from judging the same apparatus on the second days may be another way to cut down on this memory bias.

Summary. Research on gymnastics judging have investigated three biases: 1) international bias, 2) within-team order bias and 3) warm-up or memory bias. These studies have introduced the concept that the gymnasts are not always being evaluated for their performance but rather for things that are beyond their control.

Suggestions have been made to attempt to eliminate these biases. To control for international bias, scoring procedures are being adjusted so that six judges are on a panel. The gymnast's score would be the average of the middle four judges' scores.

Other bias correction factors are presently being examined. It was suggested that coaches place their team members randomly within an order to eliminate the within-team order effect (Landers, 1970). To attempt to eliminate memory biases, judges should not be present for the gymnasts' warm-up session. Gymnasts are also advised to perform their best throughout their warm-up and their competitive performance.

Potential Body Shape Bias

There may be another bias included in those already discussed. The studies presented on the body composition of gymnasts and the differences found between gymnasts and other athletes do not explain why these discrepancies occur. It could be because of the physical demands of the sport. Vercruyssen (1984) comments that there are biomechanical advantages of a short slim physique in rotational movements, arm support, and inverted balancing activities.

Another reason for the gymnasts to be short and slim could originate from the aesthetic aspect of gymnastics and how the gymnast looks when performing. The essence of an aesthetic sport such as gymnastics (and also figure skating and diving) may create a judges' bias for body shape. Whether the gymnast is tall, short, stocky or slender could have an effect on the scores that the judge gives the gymnasts' routines. Based on anecdotal reports (Ryan, 1995), this bias would come from the expectation that gymnasts should be small in stature with a low percentage body fat, not only for improved performance but also for improved appearance while performing. The only study found that begins to investigate the possibility of this bias is the study by Falls and Humphrey (1978) presented earlier. Their findings concluded that certain body

types may be a factor in women's gymnastics championships. Although their research was not completely conclusive, it does raise further questions as to whether the difference between placers and non-placers is solely skill.

Judges' Experience

Another factor that may have an impact on the judging process is the previous experience of the judge. Numerous studies have investigated the differences between novice and expert athletes and judges in the realm of physical education and sport (Armstrong & Hoffman, 1976; Bard et. al., 1980; Franks, 1993; Imwold & Hoffman, 1983; Starkes & Deakin, 1980; Ste-Marie & Lee; Thomas, 1994). For example, Armstrong and Hoffman (1979) investigated whether experienced tennis teachers could be differentiated from inexperienced tennis teachers on the basis of their ability to identify common performance errors of the tennis forehand . Forty experienced tennis coaches and 40 inexperienced undergraduate students took part in the study. The researchers found that the experienced subjects were more skilled at detecting errors than the inexperienced subjects.

Gymnastics studies. Ste-Marie & Lee (1991) also found this to be true when they investigated whether specific prior processing effects would extend to the judgment of gymnastics moves and whether these effects were dependent upon the judge's expertise. It was expected that expert judges would be better at detecting errors than novice judges. As well, they questioned whether expert judges were more or less susceptible to prior processing effects than novice judges. The researchers found that expert judges were better than novices at identifying form error, which was

consistent with earlier studies. Moves with the same performance in both phases were judged most accurately by the experienced judges, new moves less accurately, and the lowest level of accuracy occurred when the moves had differed between the two phases. Perceptual judgments were the same no matter what the judge's expertise, and these judgments were subject to specific memory influences.

In Vickers (1986) work, the same was true for experienced and inexperienced athletes. This research used a resequencing task on gymnasts with differing abilities. Their task was to reconstruct sets of photographs of gymnastics moves as quickly and as accurately as possible with time and error measures recorded. When the performance of the elite subjects was compared to that of the intermediate and novices, the elite subjects were significantly faster and more accurate than the others.

Bard et. al. (1980) analyzed the visual search patterns of four expert and three novice gymnastics judges. They found that expert judges had 27% fewer fixations than novice judges. The experts fixated more on the upper part of the gymnast's body, while the novices concentrated more on the legs. Also, it was found that novice judges detected only half the errors detected by expert judges.

Franks (1993) considered the observational accuracy of expert and novice gymnastics coaches when detecting differences between two separate technical performances of a front handspring seen on videotape. No main effects were found, but there was a significant interaction between experience and stimulus type (perfect, error). When the two performances were the same, the experts made more mistakes than the novice coaches. However, when the performances were different, experts

were superior at detecting a difference. Franks suggests that both training and job experience cause a sensitization or predisposition to view a situation in a particular way, even when alternative perceptions are possible.

Stephenson and Jackson (1977) also found the above statement to be true when they studied the effects of training and position on judges' ratings of a gymnastics event. Two groups of women received written training, and two other groups received extensive practical training. Afterwards, they judged squat vaults from one of two angles of viewing: from the side or the front of the vault. The authors found that the judges' viewing position and training had significant effects upon their ratings. Judges that viewed the vault from the front and had received extensive practical training had significantly lower ratings. The results showed that the difference in scoring may have depended upon the amount and the type of training that the judges received.

Summary. This literature suggests that experience as an athlete, coach or judge has an effect on performance, error detection, and perceptual accuracy. Type of training of the judge is a factor important to consider when analyzing effects of experience (Stephenson & Jackson, 1977). There also appears to be significant differences between experienced and novice judges and coaches when it comes to error detection (Franks, 1993; Ste. Marie & Lee, 1991). As Franks (1993) puts it, experts are predisposed to "seek out" errors in the performances they are observing. The researchers suggest that developing training programs designed to improve the perceptual abilities of observers should be considered as one method of improving the

objectivity and validity of coaching and judging evaluation (Franks, 1993; Imwold & Hoffman, 1983; Stephenson & Jackson, 1977).

Summary of Chapter

The body of the elite gymnast has changed over the years. Age, height and weight have all decreased (Bale, 1994; Bale & Goodway, 1990; Bernadot & Czerwinski, 1991; Claessens & Veer, 1991; Novak et. al., 1977; Sinning & Wayne, 1978; Thorland & Johnson, 1981). To satisfy this new definition of the gymnast's body, gymnasts have sometimes engaged in pathogenic weight control methods that are dangerous to their health (Harris & Greco, 1990; Petrie & Stoever, 1993; Rosen & Hough, 1988; Warren et. al., 1990; Sundgot-Borgen, 1994; Thorton, 1990). Studies have determined that gymnasts with varying levels of experience and age frequently do not have healthy diets for the amount of energy expenditure (Calabrese, 1985; Ersoy, 1991; Benson et. al., 1990; Lindboe & Slettebo, 1984; Reggiani et. al., 1989). When asked, gymnasts feel that they are pressured by their coaches, the nature of the sport and the subjectivity of the judging to be thin.

Studies have shown some of the biases involved in the subjectivity of evaluating performance (Ansorge et. al., 1975; Ansorge & Sheer, 1988; Sheer & Ansorge, 1975; Ste-Marie 1996; Ste-Marie & Lee, 1991; Ste-Marie & Valiquette, 1996). Psychological factors as well as the judge's expertise are factors that could help explain some of the ambiguity of evaluating gymnastics performances (Sheer & Ansorge, 1979; Ste-Marie & Lee, 1991). Only one study has begun to explain the possible reasons behind the gymnast's tendency to use dangerous eating patterns to stay small (Falls & Humphrey, 1978). More research is needed to understand fully the

dangers of subjective evaluation at all levels of gymnastics.

CHAPTER 3

Methodology

The purpose of this study was to discover whether the gymnast's body shape affects the judging of gymnastics. The goal was to test whether the judges' perceptions of the aesthetic appeal of the gymnast's body were related to the performance scores and to the body mass index of the gymnast. A second goal was to test whether the judges' experience affected their perceptions of the gymnast's body shape and the gymnasts' performance scores.

Participants

Gymnasts. Eighteen provincial level female gymnasts volunteered as participants for the development of the videotaped performances to be judged. These gymnasts were all members of Montreal area clubs and competed with other clubs in their region of Quebec. They were competitive athletes between 9 and 12 years of age, and all practiced between 12 and 15 hours a week. They performed compulsory routines on the bars, beam and floor at their level. In provincial gymnastics there are four levels: A,B,C,D, with A being the highest. The gymnasts in this study competed at the D level of competition. They were chosen because they represented a competitive regional level of gymnastics, but not an elite national level. The gymnasts were athletes that enjoyed the sport in a competitive way as a medium to learn and have fun.

Judges. Eighteen female gymnastics judges certified by the Canadian Gymnastics Federation and residing in the Eastern Region of Ontario served as subjects for this study. Nine of the judges were considered experienced, having judged for more than 10 years, and nine were novices who had been judging for less than five years. Certification levels of the judges ranged from level I to Level VI. A questionnaire was given to each judge to obtain information on her background in gymnastics. This information was used in the discussion of the results but not in the data analysis.

Instrumentation

Eighteen gymnasts were videotaped in their competition leotards at their respective gymnastics clubs. Each gymnast performed level D compulsory routines on bars, beam and floor. Vaulting was excluded from the study due to the differences in the nature of judging this event. The movements on the vault are very fast (approximately two seconds) and discrete, requiring the judges to attempt to remember what they see. It was felt that this event might produce varied scores for different reasons than the other events.

Videotaping. All videotaping was done with a color Sony video camera located at a constant position, approximately where a judge would view routines in the competition setting (see Appendix A). The positioning was consistent for all gymnastics clubs with no zooming or panning involved for the bars and beam. For these events, the apparatus filled the video screen as the gymnast performed. For the

floor routines, panning and zooming were necessary to keep the gymnast close and focused similarly to the way a judge would view the gymnast in competition. Panning and zooming for the floor routines was done as consistently as possible for each floor routine. Sound was included to simulate the noise at a gymnastics competition.

All gymnasts being videotaped at each club performed on the bars and then moved as a group to the beam and then the floor. Each of the gymnasts was told to perform the routine as if she were performing at a competition. The gymnasts were told to treat the video camera as the judge, and they presented at the beginning and end of each routine.

After all the gymnasts at the club where videotaped performing on bars, beam and floor, each girl was videotaped individually for five seconds, from the anterior view, as she stood in a natural stationary position. This was done to simulate the judge's view of the gymnast just before she competes on the event and after she has finished.

Selection of routines to be judged. An initial pool of 54 compulsory routines on bars, beam and floor were obtained from the 18 gymnasts. It was assumed that a fall in the routine might affect the way the judge scored the rest of the routine, producing a bias that would confound the findings of this study. Therefore, any routines that had one or more falls were eliminated. Forty-eight remained as potential routines for the study.

It was important to ensure that the range of body shapes was represented equally in the routines used for the study. The Body Mass Index (BMI) is a ratio which

produces a proportionate value of height and weight which represented body shape for this study. At the time of the videotaping, every gymnast was measured for her height and weight. Her BMI was then calculated using the following equation:

$$\text{Body Mass Index} = \frac{\text{Mass}}{\text{Height}^2}$$

To select equal numbers of gymnasts from across the BMI continuum, these 18 BMI scores were rank-ordered and then divided into four groups, resulting in the categories shown in Table 2.

Table 2

Ranges of BMI Within Each BMI Category

	Low (n=4)	**Medium-low (n=4)**	**Medium-high (n=5)**	**High (n=5)**
BMI	14.2-15.3	15.9-16.7	16.8-18.3	18.6-20.0

For each apparatus, three routines were randomly selected from those performed by the gymnasts in each quarter of the BMI continuum. This selection resulted in 36 routines performed by 16 of the original 18 gymnasts. Not every gymnast performed every event. For example, one gymnast having a low BMI may have appeared performing a bar routine and not a floor or beam routine, while another gymnast with a low BMI may have appeared performing bar, beam and floor routines. Table 3 shows the routines on each apparatus performed by gymnasts from each BMI

category. Lead-ins and follow-up routines also appear in Table 3; they are explained in the next section.

Table 3

Routines Performed by Gymnasts from Four BMI Categories

	Gymnast	Bars	Beam	Floor	Row Totals
Lead-ins		*	*	*	3
Low BMI	1	*	*	*	
	2			*	
	3	*	*		9
	4	*	*	*	
Medium-low BMI	1	*	*	*	
	2		*	*	9
	3	*	*	*	
	4	*			
Medium-high BMI	1		*	*	
	2	*	*		
	3	*		*	9
	4	*	*	*	
High BMI	1	*		*	
	2		*		
	3	*	*	*	9
	4	*	*	*	
Follow-ups		*	*	*	3
Column totals		14	14	14	42

Note. Astericks indicate the routine performed on each apparatus.

Sequencing of routines. To avoid potential confounding effects of viewing the same routines in a certain progression, three sequential orders of the 36 routines were constructed. Each order had exactly the same routines, but they appeared in randomly

different sequences. Three additional routines from those not selected for the 36 stimuli served as lead-ins and three others served as follow-ups (see Table 3). Lead-ins were added at the beginning of the tape to accustom the judges to the type of task that they were required to do and to avoid any confounding effects of scoring the routines that appeared first. The three follow-ups were placed at the end of the 36 routines to avoid any effects of scoring when the judge was aware of the termination of the one hour experiment. The judges were told to score all routines, but the lead-in and follow-up scores were not used in the data analysis. (This procedure is similar to Ste. Marie & Lee, 1991.) The lead-ins and follow-ups remained constant over the three orders, and only the sequential order of the 36 stimuli routines changed.

Ten seconds of blank tape followed each of the 42 routines, so that the judges had time to score the routine previously viewed. A warning tone was heard one second before the next routine was shown.

Gymnasts in static position. To examine the judge's perception of the gymnast's body shape, each of the 16 gymnasts who performed the 36 stimuli routines was shown on the videotape standing in the static position. Five other gymnasts serving as lead-ins were added before the 16 gymnasts of interest, five follow-ups were added at the end, and five fillers were scattered throughout the tape. (Other gymnasts, similar in age and level of competition were videotaped in the preliminary trial work for developing the instrumentation. These standing gymnasts were used as lead-ins , follow-ups and fillers.) Lead-ins and follow-ups were used for the same reasons as those in the sequences of routines. Fillers were added to vary the range of possible

body shape perception scores and to attempt to shift the judges' attention away from the purpose of the study. The order of the gymnasts was the same for all three of the sequential orders of routines. For example, judge A saw Order 1 of the routines and then saw the gymnasts in the anterior standing position. Judge B saw Order 2 of the routines and saw the gymnasts in the anterior standing position in the same order as judge A. Following the appearance of each gymnast, five seconds of blank tape allowed the judge time to score the gymnast for her "aesthetic appeal." A warning tone was heard one second before the next gymnast was presented. For more on the construction of the videotape instrument see Appendix B.

Procedures

Six different judges, three experienced and three novice, viewed each of the three sequential orders of routines, resulting in a total of 18 judges for the study. Table 4 shows the distribution of judges across the orders. The judges viewed the videotape individually, sitting approximately one metre from the television screen.

The judge was asked to score each routine out of a maximum of 10, deducting the appropriate marks for form and technical errors in the space provided on the judging sheet (see Appendix C). This is similar to the way the judges would score an actual competition. Immediately after judging all 42 routines, each judge saw three second views of 31 young female gymnasts, including the 16 gymnasts that performed. The judge was asked to rate each girl's body shape for aesthetic appeal for gymnastics at this level. A 10-point Likert rating scale was used, with 1 indicating least aesthetically pleasing and 10 indicating most aesthetically pleasing (see Appendix D).

The verbal instructions given to the judges by the researcher appear in Appendix E. After the task was completed, the judges were asked to fill out an experience profile questionnaire to specify whether they had been involved in gymnastics as an athlete or a coach. They were also asked their club affiliation and how many years they had participated in each area as well as their years of experience as a judge (see Appendix F).

Table 4

Distribution of Judges Across Sequential Orders of Routines

Judges	**Order of Routines**			
	One	**Two**	**Three**	**Total**
Novice	3	3	3	9
Experienced	3	3	3	9
Total	6	6	6	18

Validity

Validity was considered throughout the design of the research. The procedures simulated, as closely as possible, the actual competition setting. The following points were considered;

Videotaped content being judged.

1) All gymnasts wore their competition leotards.

2) The video camera was positioned to obtain recordings from a location similar to a judge's perspective when observing a live performance.

3) All gymnasts were told to perform as if the video camera were the judge. The gymnasts treated the video camera like it was the judges' panel and presented to the camera at the beginning and end of each routine.

4) All gymnasts competed at level D of competition. In provincial gymnastics there are four levels: A,B,C,D, A being the highest.

5) Compulsory routines were used so that the routine on each of the apparatuses remained consistent for all the gymnasts. If each gymnast had her own composition of a routine, the difference in the difficulty and presentation of certain moves could result in varying scores that were not related to the focus of the study. To control for this, routines were used that had identical skills performed by each gymnast.

Judging process.

1) The judges used score sheets similar to those used in competitions (see Appendix C).

2) The total number of routines the judges scored (42) was similar to the number a judge would evaluate in a competition.

3) The judge was given a 10 second pause to record the score of the gymnast.

Pilot Study

A pilot study was performed using three novice judges to test the reliability of the instrument. Each judge was tested once and then again three days later to assess the degree to which the judges' scoring was stable from day to day.

Intraclass correlation was conducted on the judges' perception of aesthetic appeal of the gymnast's body shape and the mean performance scores. A 3 X 2 (judges by time) ANOVA with repeated measures on the last factor was conducted on the Likert scores. This produced an $\underline{R}$ =.65. A similar ANOVA was conducted on the gymnasts' performance scores which produced an $\underline{R}$ =.99. More details of the pilot study are found in Appendix B.

Data Analysis

The four independent variables were:

1) Judge's experience (2 levels: beginner and experienced),

2) Videotape orders of the routines (3 levels: order 1,2 and 3),

3) Apparatuses (3 levels: bars, beam and floor), and

4) BMI of the gymnast (4 levels: low, medium-low, medium-high and high.)

The dependent variables were the scores for:

1) Gymnast's performance and

2) Judges' perception of the gymnasts' body shape.

To assess the judges' perception of the gymnasts' body shape, a 2 X 4 (expertise of judge by BMI category) ANOVA with repeated measures on the last factor was performed on the Likert scores. To assess effects of the gymnasts' body shape on performance scores, a 2 X 3 X 3 X 4 (expertise of judge by apparatus by order by BMI category) ANOVA with repeated measures on the last factor was performed on the gymnast's performance scores. Post-hoc planned comparisons between BMI group means were conducted to locate the significantly different groups.

The level of significance for all statistical analyses was set at $p < .05$.

CHAPTER 4

Results and Discussion

The purpose of this study was to examine the judges' perception of the aesthetic appeal of the gymnasts' body shape and its influence on the performance scores given. A second purpose was to investigate the differences in perception and performance scores for two levels of judges: experienced and novice. Results of the study are presented and discussed as they relate to each hypothesis.

Judges' Perception of Gymnasts' Body Shape

It was hypothesized that the judges' perception of the aesthetic appeal of the gymnast's body shape would be higher for gymnasts who are short and light than for those who are taller and weigh more. Mean scores for judges' perception of aesthetic appeal for gymnasts in four levels of BMI are shown in Table 5. Generally, scores decreased as BMI increased. Judges' perception scores were higher for those gymnasts who were short and light than for those who were tall and heavier. An exception was the medium-low BMI group, which was perceived as slightly more aesthetically appealing than the low BMI group.

Table 5

Mean Perception Scores for Body Shape of Gymnasts with Different BMI

BMI	Judges					
	Experienced (n=9)		Novice (n=9)		Total (N=18)	
	M	SD	M	SD	M	SD
Low	6.31	0.63	7.00	0.39	6.64	0.85
Medium-low	6.83	0.63	7.19	0.97	7.01	1.08
Medium-high	6.08	0.37	6.56	0.45	6.35	0.75
High	6.31	0.43	6.33	0.57	6.25	0.95
Totals	6.26	0.87	6.77	1.03		

Note. Scores based on Likert scale: 1 = least aesthetically pleasing, 10 = most aesthetically pleasing.

Results of the 2 X 4 (judge's experience by gymnasts' BMI category) ANOVA with repeated measures on the last factor are summarized in Table 6. A significant main effect of BMI on the judges' perception scores was found, $F(3,48) = 8.53$, $p < .001$. Post- hoc planned comparisons of means indicated that the mean perception score for the medium-low BMI gymnasts was significantly different from the mean scores for the medium-high, $F(2,16) = 9.46$, $p < .01$ and high BMI gymnasts, $F(2,16) = 22.69$, $p < .001$, but not from the low BMI gymnasts. Results of planned comparisons are summarized in Table 7.

The judges showed a slight preference for the gymnasts who had a medium-low BMI over gymnasts who had a medium high or high BMI, although the effect size was

minimal. There was no significant difference between the low and medium-low BMI category means. Figure 1 illustrates the mean scores for the judges' perception of aesthetic appeal. They found gymnasts who were a little taller and bigger with a medium-low BMI at least as aesthetically pleasing as the smallest girls for this level of gymnastics. However, as predicted, the larger gymnasts were considered less aesthetically appealing.

Table 6

ANOVA Summary for Judges' Perception of Gymnasts' Body Shape for Aesthetic Appeal Scores as a Function of Judges' Experience and Gymnasts' BMI

Sources	DF	MS	F	p	w^2
Between subjects					
Experience	1	14.22	1.05	0.320	
Error	16	13.51			
Within subjects					
BMI	3	9.50	8.53	0.000	0.57
BMI X exp	3	0.65	0.58	0.630	0.58
Error	48	1.11			

Table 7

Post-hoc Planned 36 Comparisons of Mean Perception Scores of Gymnasts' Body Shape by BMI Category

Contrasts	MS	F	p	w^2
low vs. med-low	45.56	2.39	.123	
low vs. med-hi	35.56	1.67	.220	
low vs. hi	35.56	2.42	.121	
med-low vs. med-hi	69.89	9.46	.002	0.079
med-low vs. hi	93.89	22.69	.000	0.079
med-hi vs. hi	3.56	0.50	.613	

Note. df = 2,16

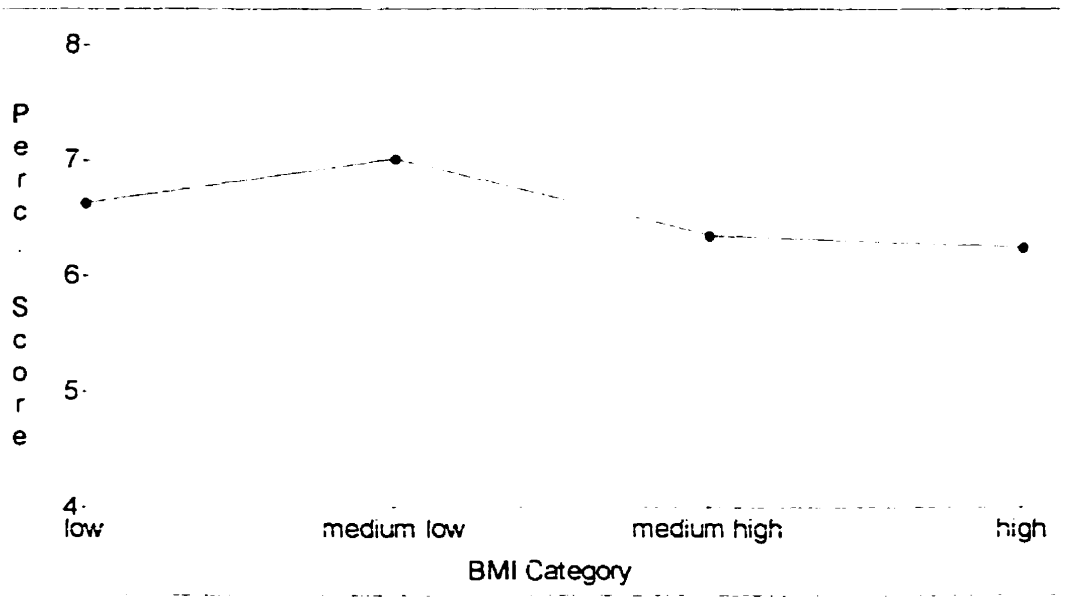


Figure 1. Judges' (N = 18) mean perception scores for aesthetic appeal of body shape for gymnasts having a low, medium-low, medium-high and high BMI.

This preference for lower BMI gymnasts was expected due to the changes observed in elite gymnasts' body shape over the last 20 years, from a bigger mature body to a smaller childlike body (Bale, 1994; Bale & Goodway, 1990; Bernadot & Czerwinski, 1991; Claessens & Veer, 1991; Novak et. al., 1977; Sinning & Wayne, 1978; Thorland & Johnson, 1981). Over these years the sport of gymnastics has changed drastically, thereby encouraging different body types. Now, elite gymnasts are younger, shorter and lighter. This produces expectations and certain preferences for a smaller more compact body shape.

Gymnasts' Performance Scores and Their Body Shape

The main purpose of this study was to discover whether the gymnast's body shape, as defined by her body mass index, influenced the scoring of her performance. It was hypothesized that the judges' scores awarded for the gymnasts' performance would be higher for gymnasts who are short and weigh less than for those who are taller and weigh more. Mean performance scores awarded by the 18 judges to gymnasts in the four levels of BMI are shown in Table 8. Generally, the gymnasts' performance scores increased as their BMI category increased.

Table 8

Mean Performance Scores for Gymnasts with Different BMI

BMI	Judges					
	Experienced (n=9)		Novice (n=9)		Total (N=18)	
	M	SD	M	SD	M	SD
Low	9.04	0.24	9.23	0.23	9.12	0.25
Medium-low	9.09	0.35	9.29	0.31	9.19	0.32
Medium-high	9.10	0.11	9.30	0.12	9.20	0.12
High	9.30	0.22	9.50	0.22	9.41	0.21
Totals	9.16	0.24	9.35	0.24		

The initial 2 X 3 X 3 X 4 (judges' experience by order by apparatus by gymnasts' BMI category) ANOVA with repeated measures on the last factor found no significant differences between performance scores as a result of the apparatus ($p >$

.05). Therefore, the scores were collapsed to a mean performance score for each gymnast over all apparatuses. The apparatus variable was eliminated from subsequent analyses.

Results of the 2 X 3 X 4 (judges' expertise by order by gymnast's BMI) ANOVA with repeated measures on the last factor are summarized in Table 9. A significant main effect of BMI on the performance scores was found $F(3,36) = 44.7$, $p < .001$. Generally, the analysis showed that gymnast's performance scores increased as their BMI increased (see Figure 2). Post-hoc planned comparisons indicated that the performances of the gymnasts having a high BMI were judged significantly higher than those of all the other gymnasts. Table 10 summarizes the results of the post-hoc planned comparisons of mean performance score by BMI category.

Table 9

ANOVA Summary for Gymnasts' Performance Scores as a Function of the Order of Presentation of the Routines and the Judges' Experience

Source	DF	MS	F	p	w^2
Between subjects					
Exp	1	2.82	4.35	.059	0.34
Order	2	2.41	3.71	.056	0.42
Exp*Order	2	3.62	5.56	.020	0.30
Error	12	0.65			
Within subjects					
BMI	3	1.00	44.7	.000	0.27
BMI*Exp	3	0.002	0.08	.968	
BMI*Order	6	0.011	0.48	.822	
BMI*Exp*Order	6	0.04	1.20	.104	
Error	36	0.02			

Table 10

Post-hoc Planned Comparisons on Performance Scores by BMI Category

Contrasts (BMI)	MS	F	p	w^2
low vs. med-lo	0.34	1.46	.271	
low vs. med-hi	0.34	2.15	.122	
low vs. hi	3.65	28.43	.000	0.15
med-lo vs. med-hi	0.03	0.15	.985	
med-lo vs. hi	2.45	10.86	.000	0.14
med-hi vs. hi	2.05	18.42	.000	0.15

Note. df = 6,12

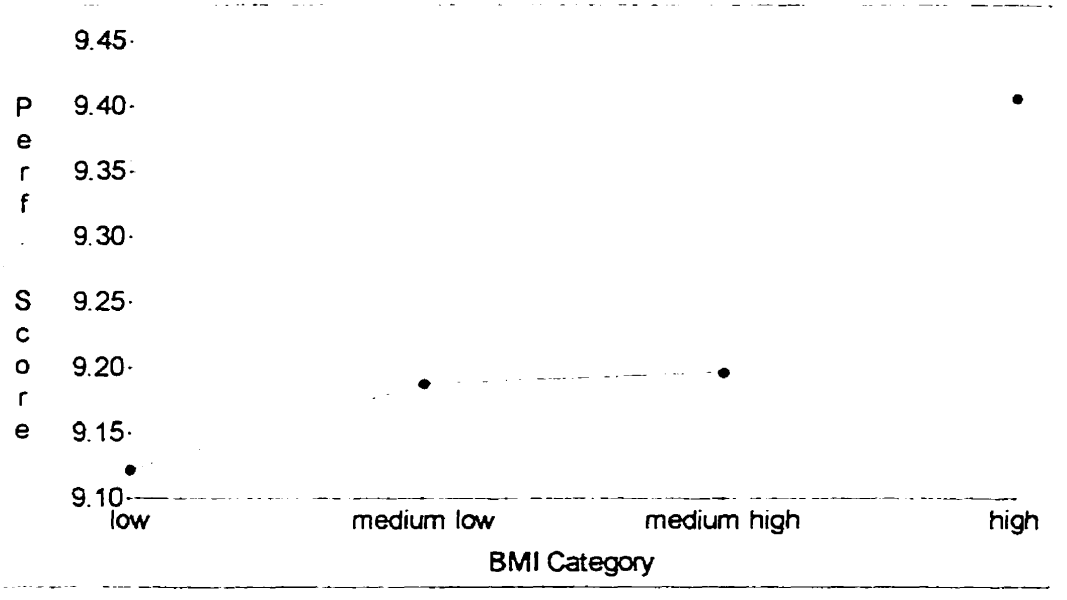


Figure 2. Mean performance scores for gymnasts with low, medium-low, medium-high and high BMI.

The findings of this study seem to support anecdotal reports (Bale, 1994; Bernadot & Czerwinski, 1991; Calabrese, 1985; Warren et. al., 1990) about the aesthetic appeal of a small gymnast's body (see Figure 1). Therefore, it was expected

that as the gymnasts' BMI scores decreased, apparently resulting in a more pleasing appearance, their performance scores would increase. Falls and Humphrey (1978) suggest this possibility for collegiate gymnasts. Their findings suggested that certain body types may be a factor in championship performance scores in women's gymnastics.

The results of the present study showed the opposite effect: as the BMI decreased, so did the gymnast's mean performance score. Initially, it was thought that the gymnast's age could be a possible explanation for the differences in performance scores. An older gymnast tends to have had more experience and may exhibit better poise and control when performing. She may be stronger, and because of extra years of practice, perhaps better skilled. A younger gymnast may not be as strong and may not have had the experience of performing routines in the competition setting. The gymnasts in this study ranged from ages 9 to 12. Table 11 shows the distribution of ages for the 16 gymnasts in the BMI divisions. An analysis of variance with age as the repeated measures factor showed no significant difference of mean performance scores across age groups ($p > .05$). Therefore, it was concluded that age itself was not the critical factor.

The results did not support the major hypothesis which stated that judges would award higher performance scores to gymnasts with a low BMI than for gymnasts with a high BMI. It is possible that the scores might reflect physical strength of the gymnast since scores increased with BMI. It is suspected that at this level of gymnastics, the gymnasts with a higher BMI simply demonstrated more strength, enabling them to

perform better. A logical assumption seems to be that the judges rewarded skill, which perhaps reflected the "larger" gymnasts' greater strength.

Table 11

Distribution of Gymnasts' Age Across the BMI Categories

BMI	Low	Medium-low	Medium-high	High
Ages	9, 9, 10, 12	10, 10, 11, 11	10, 11, 12 ,12	11, 11 , 11 ,12

Judges' Experience

It was hypothesized that the judge's experience would affect both the way she viewed the gymnasts' body shape for aesthetic appeal and the scores awarded for the gymnastics performances.

Perception scores . The analysis of variance showed that there was no significant difference between the gymnast's aesthetic appeal scores for the two levels of judges (see Table 6).

It had been expected that experienced judges would score gymnasts lower for aesthetic appeal than would the novice judges. This was expected for two reasons: experienced judges have had more years of experience evaluating performance, and therefore have had time to develop a stronger perception of their preferred body shape. Novice judges have not been evaluating performance for very long and may not have as strong a perception of an ideal body shape. Secondly, the experienced judge has had more opportunity to evaluate elite levels of gymnastics where the gymnasts tend to be small (Bernadot & Czerwinski, 1991; Claessens et. al., 1992).. Although experienced

judges seemed to prefer smaller gymnasts' body shape more than the novice judges did, the difference was not significant. This finding might be explained by the judges' experience profiles and Fyock and Stangor's (1994) stereotype maintenance theory.

According to Fyock and Stangor (1994), stereotypes are self-maintained because people tend to remember expectancy-confirming information better than expectancy-disconfirming information. For example, a person involved in the sport of gymnastics for many years has worked with and evaluated, consciously or unconsciously, many gymnasts with varying body shapes and ability. A person who expects smaller gymnasts to be more skilled will tend to remember information that confirms this better than conflicting information. Skilled gymnasts who are bigger will not be recalled as clearly.

In this study, it was expected that an experienced judge would have a better maintained expectation of gymnasts' body shape. All judges showed that they preferred the smaller gymnasts' body over the bigger gymnasts' bodies. The judges' experience profiles indicated that almost all the judges had also been athletes and/or coaches as well as judges (see Table 12). These experiences could have affected what the judges consider to be an aesthetically pleasing body shape (Franks, 1993; Stephenson & Jackson, 1977). Since all judges had extensive experience in the domain of gymnastics, they may all have had a well developed sense of preferred body shape. Having been involved with the sport of gymnastics in other capacities may have played a part in reinforcing these judges' development of their preference for a smaller build of gymnast.

Table 12

Judges' Years of Experience as a Judge, Athlete and Coach

	Years of Experience					
	Judge		Athlete		Coach	
Judge Level	M	SD	M	SD	M	SD
Experienced	14.33	5.20	5.22	3.80	10.89	6.77
Novice	2.57	1.24	11.00	7.21	14.78	10.99

Performance scores. It was hypothesized that experienced judges would award lower performance scores than novice judges. The main effect for experience was very close to being significant $F(1,12) = 4.345$, $p < .06$ (see Table 9). The means for experienced judges and novice judges (see Table 8) show that the experienced judges did award lower scores ($M = 9.16$) than the novice judges ($M = 9.35$).

The judge's task in this study was to deduct marks for form errors. The difference between the performance scores of experienced and novice judges suggests that the experienced judges were somewhat better at detecting errors than novice judges. Ste-Marie & Lee (1991) found this to be true when they investigated whether specific prior processing effects would extend to the judgment of gymnastics moves and whether these effects were dependent upon the judge's expertise. Expert judges were better than novices at identifying form errors. These findings are also consistent

with those of Allard and Starkes (1980), Armstrong and Hoffman (1979), and Biscan and Hoffman, (1976).

A significant interaction between judge's experience and the sequential order of the routines was found, $F(2,12) = 5.56$, $p < .05$ (see Table 9). The interaction is illustrated in Figure 3.

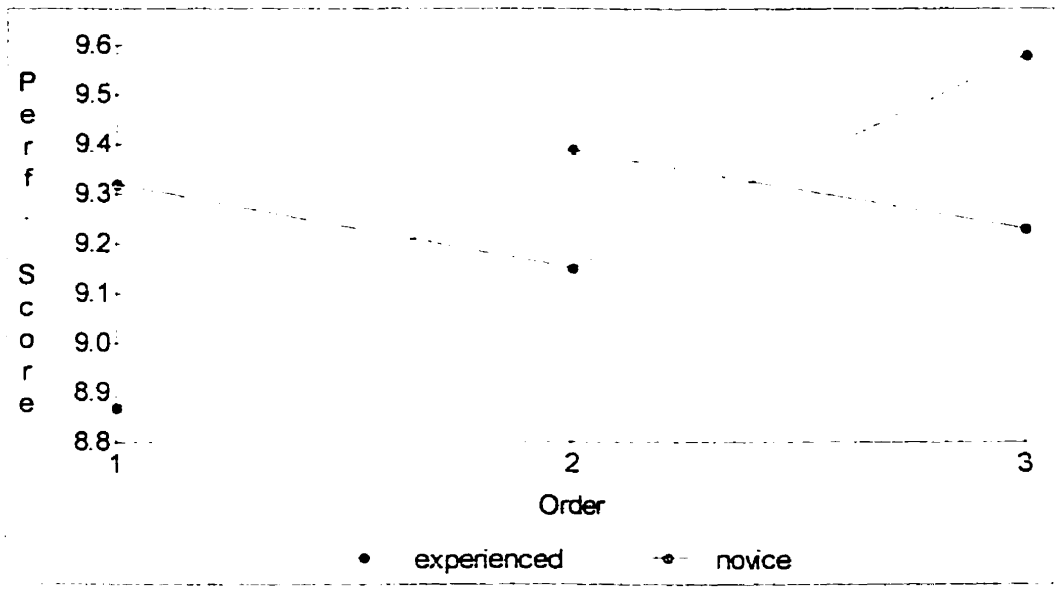


Figure 3. Interaction of judge expertise and sequential order for gymnasts' performance scores.

Three different orders of routine presentation were assembled to control for the possibility of a confounding effect due to order. The difference in mean scores for the orders ($M = 9.10, 9.27, 9.41$, for Orders 1,2,3 respectively) may be due to the random placement of judges within each order. A judge who scores substantially higher or lower in one order can either increase or decrease the mean score. It was expected that experienced judges would award lower scores overall than novice judges. For Orders 1

and 3 this hypothesis holds true. However, in the second order, two of the highest means are generated by experienced judges (see Table 11). Both means (Ms = 9.53) are quite a bit above the overall mean of the experienced judges (M = 9.16), causing the experienced judge mean for this order to exceed that of the novice judges. In addition, in Order 3, one novice judge gave a mean score of 9.79 which was considerably above the overall novice judge mean of 9.35 (see Table 13). This could be due to the fact that this judge is very lenient and generally gives higher scores. When a judge is on a panel for a competition, they adjust their scores to match the range of the other judges after the first few routines. The judges could not do that in this case and as a result her scores are at the higher end of the range. In Order 1 another novice judge gave a mean score of 8.93, which was far below the mean score of the others (M =9.35). These variances resulted in a significant effect of the interaction of the experience and order factors on the performance scores but a non-significant main effect for the factors. The differences in scores also demonstrate the lack of consistency between the judges.

Table 13

Mean Performance Scores Within Each Order for Experienced and Novice Judges

		Judges				
Order		**Experienced**		**Novice**		
		M	**SD**	**M**	**SD**	**Order Means**
		8.80	0.29	9.54	0.28	
1		8.87	0.17	9.42	0.27	
		8.90	0.38	8.93	0.41	
	M	8.84	0.28	9.29	0.32	9.10
		9.53	0.26	9.22	0.24	
2		9.05	0.34	9.00	0.35	
		9.53	0.14	9.14	0.29	
	M	9.37	0.25	9.12	0.30	9.27
		9.28	0.34	9.79	0.15	
3		9.17	0.31	9.57	0.15	
		9.11	0.35	9.37	0.24	
	M	9.19	0.33	9.58	0.19	9.41
Overall M		9.16	0.24	9.35	0.24	

Summary of Chapter

As hypothesized, the judges' perception of the aesthetic appeal of the gymnasts' body shape generally increased as gymnasts' BMI decreased. That is, these judges preferred the body shape of gymnasts in the lower half of the range of BMI seen in this study. This finding supports previous anecdotal reports (Bale, 1994; Bernadot & Czerwinski, 1991; Calabrese, 1985; Warren et. al., 1990). Although the mean

perception scores for experienced judges were lower than those of the novice judges, the difference was not significant. The judges' experience profiles may help explain this finding. Having been involved for a considerable period of time with the sport of gymnastics in capacities other than just judging, these judges may all have had a well developed perception of a preferred gymnast's body shape.

The results showed a significant effect of BMI category on performance scores. Judges scored those gymnasts in the highest BMI category significantly higher for performance than all other gymnasts, and performance scores generally increased as the gymnasts' BMI category increased. BMI represented body shape in this study and the judges preferred the aesthetic appeal of lower BMI (smaller) gymnasts. Therefore, the increasing performance scores with BMI was contradictory to the hypothesis of the study. Apparently, these judges were not affected by their personal preference for body shape when evaluating performance. The results also suggest that judging experience may facilitate the detection of errors in performance.

Overall the judges found gymnasts with a low to medium-low BMI most aesthetically pleasing. This preference for a smaller body shape did not influence the gymnasts' performance scores. In fact, the trend for the performance scores and BMI category was the reverse of the trend for perception scores and BMI category.

Chapter 5

Summary and Conclusions

The purpose of this study was to determine whether judges give higher scores to those gymnasts at the club level of competition who are shorter and weigh less. The study attempted to clarify whether the judges' perception of the aesthetic appeal of the gymnasts' body shapes increases as the gymnasts' BMI decreases. A second concern was whether gymnastics judges show evidence of a bias for a particular body shape in the scores they award for performance. Finally, the study considered whether perception and performance scores differ for experienced judges (>10 years experience) in comparison with novice judges (<5 years experience).

Methodology

Sixteen female gymnasts between 9 and 12 years of age, at the D level of competition performed compulsory gymnastics routines on bar, beam and floor. All routines were videotaped. Height and weight measurements were taken of the 16 gymnasts in order to calculate their Body Mass Index. Thirty-six of the videotaped routines were randomly selected, with the conditions that there were equal numbers of gymnasts from each BMI category and the three apparatuses were also equally represented. These routines were arranged in three different random orders for

viewing by the judges. In addition, each of the 16 gymnasts was shown on videotape standing in the static position.

Eighteen judges, nine experienced and nine novice, watched and scored the videotaped compulsory gymnastics routines. Then they observed the gymnasts standing in the static position and rated their perception of the aesthetic appeal of the gymnasts' body shape for this level of gymnastics. Perception and performance scores were analyzed using repeated measures analyses of variance.

Summary of Findings

Hypothesis one considered the judges' perceptions of aesthetic appeal of the gymnasts' body shape. As expected they preferred the body shape of those gymnasts with a low to medium-low BMI over those gymnasts with a medium-high and high BMI. It was also hypothesized that the experience of the judge would have an effect on the perception scores. The experienced judges were expected to score all gymnasts lower for aesthetic appeal than the novice judges. Although this was found to be true, the difference in scores was not significant.

The results did not support the second hypotheses. It had been predicted that judges would give higher performance scores to those gymnasts who were shorter and weighed less (low BMI) than to those who were taller and weighed more (high BMI). The results showed an opposite trend. Judges tended to award higher scores to those gymnasts with a higher BMI than to those with a low BMI. It was thought that age of the gymnasts could have been a factor in this trend. However, after grouping the scores by age and conducting another analysis of variance, the findings showed that

age was not a significant factor. In summary, judges preferred to see a smaller more compact gymnast, but this expectation did not bias the scores they awarded for performance.

The effect of the experience of the judge on the routine scores awarded approached significance. Experienced judges tended to award lower scores than did novice judges. Since the judges' task was to deduct marks for form errors, it can be said that experienced judges detected errors in performance better than novice judges. Literature supports the conclusion that experienced judges have a better eye for detecting performance errors than novice judges (Allard & Starkes, 1980; Armstrong & Hoffman, 1979; Biscan & Hoffman, 1976; Ste. Marie & Lee, 1991).

Conclusions

Based on the results of this study, and within the limits of the study, the following conclusions can be made:

1. Judges find gymnasts having a smaller body more aesthetically pleasing.

2. The judges' perception of the aesthetic appeal of a gymnast's body shape may not affect the scores they award for the gymnast's performance.

3. Bigger gymnasts may perform compulsory routines better than smaller gymnasts.

4. The judge's experience in gymnastics judging may not affect her perception of the aesthetic appeal of the gymnast's body shape.

5. The judge's experience in gymnastics judging may or may not affect the performance scores she awards.

Implications

Judges like to see gymnasts that are small; however, this preference need not affect the performance scores awarded. In this study, there appeared to be no bias for a smaller body shape. In fact, the performance scores showed a strong trend favoring the gymnasts who were larger.

It may be that the bigger girls (in the high BMI category) demonstrated more strength and control than the smaller gymnasts and scores were awarded accordingly. For this level of gymnastics, athletes and coaches should be aware of the differences in strength and control of the competitors when attempting to justify performance scores. It is easy to justify performance scores when a gymnast has fallen or committed a serious error in her routine. However, when all the gymnasts have performed similarly, it is best to trust that the judges have deducted the appropriate points for smaller form errors and have considered the poise of the gymnast and the presentation of the skills (extension of arms and legs and clean, finished movements). In addition, the coaches should recognize the importance of performing routines with poise and composure and should relay this to their athletes. This could be the difference between a placer and a nonplacer in a gymnastics competition.

The judges should be reminded of the level of competition they are judging. This should be emphasized so that they can consciously keep in mind that any

previously formed expectations (e.g., of elite gymnasts) should not influence the way in which they score the gymnasts' performances.

The results of this study found that judges' experience had a close to significant effect on performance scores. It may be beneficial to have novice and experienced judges make up a panel at both low and high levels of competition. When the scores are averaged, expectations that might influence the scores can be filtered out. For example, it is possible that the experienced judges are better at detecting form errors and, therefore, produce lower performance scores than the novice judges. A panel of all experienced judges at a low level of competition might produce lower scores than are normally expected. This can discourage gymnasts. On the other hand, a panel of all novice judges may produce higher scores that do not effectively reflect errors in the performance.

The possible effects of subjective judging should be addressed throughout the judges' training. Coaches and athletes must understand possible evaluation biases which could cause score fluctuation. It is also important to be sure that coaches do not pressure gymnasts to attempt to control their body shape.

There is a need for valid research to continue the investigation into biases in judging. This study has attempted to isolate the potential body shape bias. A trend was found that was opposite to what was expected. Even the researcher's expectation is a bias. Therefore, extensive research is needed to justify or counter society's beliefs.

Recommendations for Further Study

Researchers have only begun to investigate the problems of subjective evaluation. Specifically, it is still not entirely clear as to how body shape affects gymnastics judging. On the basis of results of this present study, the following studies are recommended:

1. A similar study using an elite level of gymnasts. Compare the results with the present results and those of Falls and Humphrey (1978).

2. A similar study to establish the effects of the gymnasts' body shape on aspects other than error detection, such as the rhythm and tempo and style components.

3. A deeper qualitative look into why judges perceive gymnasts with a lower BMI more pleasing to the eye than a gymnast with a high BMI.

4. A cross-cultural investigation of the judges' reasoning in terms of how judges from different culture perceive gymnasts with a lower BMI.

5. An investigation of the social psychological factors of the judge and their effect on the judges' perception of aesthetic appeal of the gymnasts' body

shape and the evaluation process. For example, answer the question: how does the judges' locus of control affect their scores for performances of gymnasts with differing body mass indices?

6. Further investigation into the effects of the judges' experience on judging gymnastics. More specifically, responses to bias when judges have had no background in gymnastics as an athlete or a coach in comparison to those that have had extensive experience as a gymnast or a coach.

References

Ansorge, C. J., & Scheer, J.K. (1988). International bias detected in judging gymnastic competition at the 1984 Olympic Games. Research Quarterly for Exercise and Sport, 59, 103-107.

Ansorge, C.J., Scheer. J.K., Laub, J., & Howard, J. (1978). Bias in judging women's gymnastics induced by expectations of within team order. Research Quarterly, 49, 399-405.

Armstrong, C.W., & Hoffman, S.J. (1979). The effects of teaching experience , knowledge of performer competence and knowledge of performance outcome on performance error identification. Research Quarterly, 50(3), 318-327.

Bale, P. (1994). Body composition and menstrual irregularities of female athletes. Sports Medicine, 17(6), 347-352.

Bale, P. & Goodway, J. (1990). Performance variables associated with the competitive gymnast. Sports Medicine, 10(3), 139-145.

Bard, C., Fleury, M., Carriere, L., & Halle, M. (1980). Analysis of gymnastics judges visual search. Research Quarterly for Exercise and Sport, 51(2), 267-273.

Benson, J.E., Allemann, Y., Theintz, G.E., & Howald, H. (1990). Eating Problems and Calorie Intake Levels in Swiss Adolescent Athletes. International Journal of Sports Medicine, 11, 249-252.

Bernadot, D., & Czerwinski, C. (1991). Selected body composition and growth measures of junior elite gymnasts. Journal of the American Dietetic Association, 91(1), 29-33.

Beunen, G., Claessens, A., & Van Esser, M. (1981). Somatic and motor characteristics of female gymnasts. In Borms, J. et. al. (Eds) The Female Athlete, Medicine and sport XV, 170-185, Karger, Basel.

Brawley, L.R., Landers, D., Miller, L., & Kearns, K. (1979). Sex bias in evaluating motor performance. Journal of Sport Psychology, 1, 15-24.

Bruno, J.F. (1986). Scoring bias in springboard and platform diving : A bayesian approach to revision of DD tables. Research Quarterly for Exercise and Sport, 57(1), 16-22.

Calabrese, L.H. (1985). Nutritional and medical aspects of gymnastics. Clinics in Sports Medicine, 4(1), 23-30.

Caladrone, G., Legise, M., Giametro, M., & Berlutti, G. (1986). Anthropometric measurements, body composition, biological maturation and growth predictions in young female gymnasts of high agonistic level. Journal of Sports Medicine and Physical Fitness, 26, 263-273.

Claessens, A.L., Malina, R.M., Lefevre, J., Beunen, G., Stijnen, V., Maes, H., & Veer, F.M. (1992). Growth and menarcheal status of elite female gymnasts. Medicine and Science in Sports and Exercise, 24(7), 755-763.

Claessens, A.L., Veer, F.M., Stijnen, V., Lefevre, J. Maes, H., Steens, G., & Beunen, G. (1991). Anthropometric characteristics of outstanding male and female gymnasts. Journal of Sport Science, 9, 53-74.

Die, A.H, & Holt, R. (1989). Perceptions of the typical female male, female athlete and male athlete. International Journal of Sport Psychology, 20, 135-146.

Ersoy, G. (1991). Dietary status and anthropometric assessment of child gymnasts. The Journal of Sports Medicine and Physical Fitness, 31(4), 577-580.

Falls, H.B., & Humphrey, D. (1978). Body type and composition differences between placers and nonplacers in an AIAW gymnastics meet. Research Quarterly, 49, 39-43.

Franks, I.M. (1993) The effects of experience on the detection and location of performance differences in a gymnastic technique. Research Quarterly for Exercise and Sport, 64(2), 227-231.

Fyock, J., & Stangor, C. (1994). The role of memory biases in stereotype maintenance. The British Journal of Social Psychology, 33, 331-343.

Gealdelmann, P.L. (1980). Physical education: A stronghold of sex role stereotyping. Quest, 32(2), 192-200.

Guise, J. (1996, July). Inshape: Olympic Highlights. Shape, 15(11), 30.

Harris, M.B., & Greco, D. (1990). Weight control and weight concern in competitive female gymnasts. Journal of Sport & Exercise Psychology, 12, 427-433.

Imwold, C.H., & Hoffman, S.J. (1983). Visual recognition of a gymnastics skill by experienced and inexperienced instructors. Research Quarterly for Exercise and Sport, 54(2), 149-155.

Lindboe, C.F., & Slettebo, M. (1984). Are young female gymnasts malnourished? European Journal of Applied Physiology, 52, 457-462.

Loosli et. al. (1986). Nutrition habits and knowledge in competitive adolescent female gymnasts. The Physician and Sportsmedicine, 14(8), 118-138.

Novak, L.P., Woodward, W.A., Bestit, C., & Mellerowicz, H. (1977). Working capacity, body composition and anthropometry of Olympic female athletes. Journal of Sports Medicine, 17, 275-283.

Petrie, T., & Stoever, S. (1993). The incidence of bulimia nervosa and pathogenic weight control behaviors in female collegiate gymnast. Research Quarterly for Exercise and Sport, 64(2), 238-241.

Reggiani, E., Arras, G.B., Trabacca, S., Senarega, D., Chiodini, G. (1989). Nutritional status and body composition of adolescent female gymnasts. The Journal of Sports Medicine and Physical Fitness, 29(3), 285-288.

Rosen, L.W., & Hough, D.O. (1988). Pathogenic weight-control behaviors of female college gymnasts. The Physician and Sportsmedicine, 16(9), 141-144.

Ryan, J. (1995). Little Girls in Pretty Boxes. New York, NY: Doubleday.

Ryan, J. (1996, April). The years of training dangerously. Shape, 15(10), 121-123, 164-166.

Scheer, J.K., & Ansorge, C.J. (1975). Effects of naturally induced judges' expectations on the ratings of physical performances. Research Quarterly, 46, 463-470.

Scheer, J.K., & Ansorge, C.J. (1979). Influence due to expectations of judges: A function of internal-external locus of control . Journal of Sport Psychology, 1, 53-58.

Scheer, J.K., Ansorge, C.J., & Howard, J. (1983). Judging bias induced by viewed contrived videotapes: A function of selected psychological variables. Journal of Sport Psychology, 5, 427-437.

Starkes, J.L., Deaken, J. (1984). Perception in sport: A cognitive approach to skilled performance. In W.F., Straub & J.M. Williams (Eds), Cognitive Sport Psychology, 115-128. Lansing, N.Y.: Sport Science Associates.

Ste-Marie, D.M. (1996). International bias in gymnastic judging: Conscious or unconscious influences? Perceptual and Motor Skills, 83, 963-975.

Ste-Marie, D.M., & Lee, T.D. (1991). Prior processing effects on gymnastic judging. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 126-136.

Ste-Marie, D.M., & Valiquette, S.M. (1996). Enduring memory-influenced biases in gymnastics judging. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(6), 1498-1502.

Stephenson, D.A., & Jackson, A.S. (1977). The effects of training and position on judges' rating of a gymnastic event. Research Quarterly, 48(1), 177-180.

Sundgot-Borgen, J. (1994). Risk and trigger factors for the development of eating disorders in female elite athletes. Medicine and Science in Sports and Exercise, 4, 414-419.

Thomas, K., (1994). The development of sport expertise: From Leeds to MVP Legend. Quest, 46, 199-210.

Thomas, K., & Thomas, J.R. (1994). Developing expertise in sport: The relation of knowledge and performance. International Journal of Sport Psychology, 25, 295-312.

Thorland, W.G., Johnson, G.O., Fagot, T.G., Tharp, G.D, & Hanner, R.W. (1981). Body composition and somatotype characteristics of junior Olympic athletes. Medicine and Science in Sports and Exercise, 13(5), 332-338.

Thornton, J.S. (1990). Feast or famine: Eating disorders in athletes. The Physician and Sportsmedicine, 18(4), 116-122.

Tveit, P. (1989). Can dieting make you fat? International Gymnast, May, 55.

Vickers, J. (1986). The resequencing task: Determining expert-novice differences in organization of a movement sequence. Research Quarterly for Exercise and Sport, 57(3), 260-264.

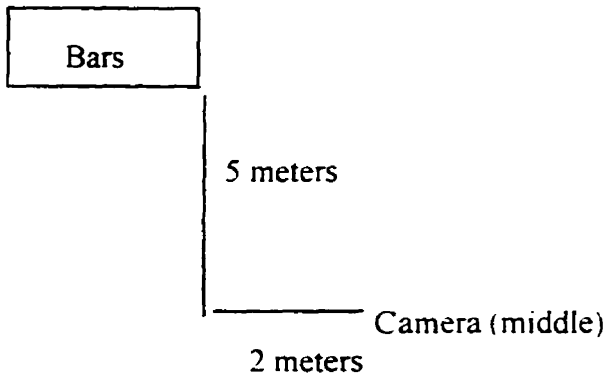
Warren, B.J., Stanton, A.L, & Blessing, D.L. (1990). Disordered eating patterns in competitive female athletes. International Journal of Eating Disorders, 9(5), 565-569.

Whissel, R., Lyons, S., Wilkinson, D., & Whissell, C. (1993). National bias in judgments of Olympic level skating. Perceptual and Motor Skills, 77, 355-358.

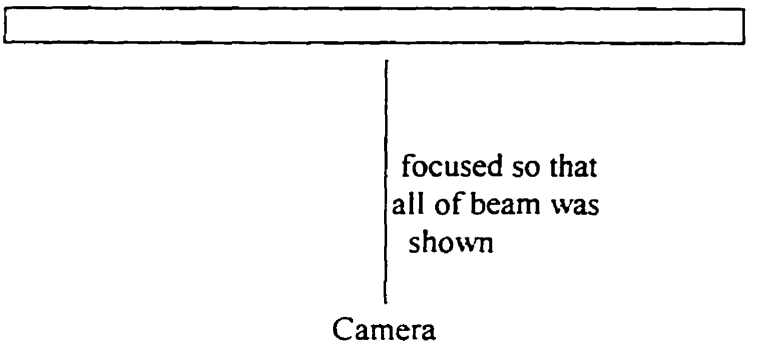
APPENDIX A

Camera Angles when Videotaping the Compulsory Routines

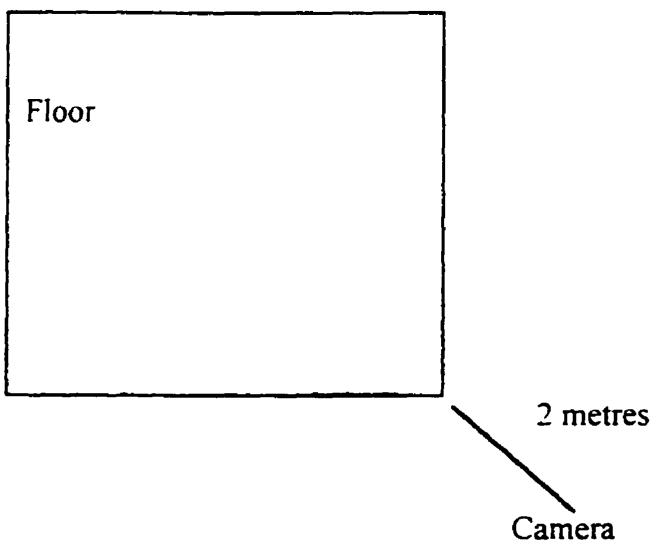
Bars



Beam



Floor



APPENDIX B

Table of Contents for Appendix B

Page

Instrumentation used in Previous Gymnastics Studies 84

Studies without videotape as instrumentation 84

Studies using videotape as instrumentation 85

Purpose 86

The Construction of the Instrument 87

Controlling for Confounding Variables 87

The gymnastic routine 87
The gymnast's skill 88
The gymnast's leotard 88
Video camera position 88
Order effects 88

Obtaining a Pool of Gymnastics Routines 88

Videotaping 89
Recording height and weight 89
Routines used for the orders 90
BMI ranking of the gymnasts 90
Organization and numbering of routines 91
Random selection of routines for the orders 91
The order of the gymnasts for judge's perception 92

Videotape Editing 92

Gymnasts' routines for the orders 92
Gymnasts for judges' perceptions 93

The Pilot Study 94

Data Collection 94
Informal questioning 94

Results 95
Correlations of the scores 95

Testing the Effects of Gymnast's Body Shape on Gymnastics Judging: Development of An Instrument

In order to develop a tool to test the effects of any factors on gymnastics judging two things can be done. The effects can be tested during the actual competition and judging setting or they can be studied in a simulated gymnastics setting reproduced as realistically as possible. In many most cases it is difficult to use the live judging setting because of the implications the study might have on the gymnasts. Videotaping the routines from a competition or from a simulated environment can be considered. The following appendix will explain the reasoning for choosing the instrumentation used for the study and how it was constructed. Also, results of the pilot study using the instrument will be displayed.

Instrumentation in Previous Gymnastics Studies

Studies without videotape as instrumentation. Research has been conducted on judging in gymnastics where the methodology did not require videotaping. These studies relied on the scores taken from the actual competition setting. Ansorge and Scheer (1980) used the complete results and judging assignments obtained from the 1984 men's and women's Olympic gymnastics team competition. They investigated whether or not a pattern of bias was shown by men's and women's gymnastics judges.

Similarly, Falls and Humphrey (1978) used scores from an Association for Intercollegiate Athletics for Women (AIAW) meet to compare performance results with body composition and somatotype. They wished to resolve whether meet placers differed from nonplacers on body composition and somatotype. Actual results were obtained, and the body measurements were taken at the competition.

Studies using videotaping as instrumentation. Little research has been done on the subjectiveness of gymnastics judging. Of the studies completed, most have relied on the use of videotaping. For example, Scheer and Ansorge (1978) used videotapes of the team optional session of the 1973 Big Eight gymnastics meet. They edited the routines to study the effects of naturally induced judges' expectations on the ratings of physical performances. They investigated the possibility that the scores of nationally certified gymnastics judges were influenced by the expected placement of gymnasts within a team order of competition. They videotaped and edited the routines so that the within-team orders of competition were reversed in one or the other of the two sets of videotapes for selected teams in each event.

Ste-Marie and Lee (1991) used videotapes of gymnast's moves to examine the influence of prior judgments on direct and indirect memory tests in gymnastics judging. They required a number of gymnastics skills performed with one form error, and then with no form errors (performed perfectly). The moves were edited into one second video clips. The purpose of presenting only a portion of the entire gymnastics

move was to standardize the length of each item and to reduce expected ceiling effects in detecting form errors.

Purpose

The purpose of the present study was to determine whether judges give higher scores to those gymnasts at the club level of competition in gymnastics who are shorter and weigh less. The study has attempted to clarify whether gymnastics judges produce different scores when they view compulsory routines by gymnasts with different body shapes. A second concern was whether the judges' perception of the aesthetic appeal of the gymnasts' body shape related to the score that they gave the gymnast for performance of her routine. Finally, the study has considered whether a bias for body shape differs for experienced judges (>10 years experience) in comparison with beginner judges (<5 years experience). Videotaping was used in order to accumulate enough compulsory routines without disrupting a competition.

Construction of the Instrument

The instrumentation that attempted to fulfill the above purpose was crucial to the success of the study. The study required a controlled number of routines from gymnasts with varying body shapes. It would be very difficult to conduct this type of study in a live competition setting. Investigating the problem would cause interruptions in the gymnastic competition, and might alter the proper judging of the gymnasts in the competition. Therefore, it was necessary to develop an instrument that would simulate the competitive setting and help determine whether the gymnast's body shape influences the gymnastics judge.

Controlling for Confounding Variables

The gymnastics routine. For this type of research, it was important to factor out as many potential confounding variables as possible. The most threatening variable was the actual routine. The routine on each of the apparatuses had to remain consistent over the gymnasts. If each gymnast had her own routine composition, the difference in skill difficulty and presentation could result in varying scores from the judges. The differences in the judges' scores could result from the variation in skills and not because of the body shape factor being studied. To control for this, routines having identical skills were videotaped. Compulsory routines (routines that have the same skills and dance movements) were the solution to having every gymnast perform the same skills. Quebec is the only province that requires gymnasts at the A,B,C, and

D level of provincial gymnastics to perform compulsory routines. Level D compulsory routines from four different clubs in the greater Montreal area were videotaped.

The gymnast's skill. A second potential confounding variable was the gymnasts' skill level. The differences in scores could be caused strictly by the variation of the gymnasts' talent if they were vastly different. For this study it was imperative that the difference in scores be due to the gymnast's body shape. Using one skill level (D) was the solution. The routines were randomly chosen and then edited.

The gymnast's leotard. In order to simulate the competition setting and avoid confounding effects of various leotards, the gymnasts were required to wear their competition leotard.

Video camera position. The video camera was used with a tripod and placed at a consistent location for the taping of all the routines at the different clubs. The routines were videotaped from a position similar to that from which a judge would view the routines in a competition. Appendix A is a diagram showing the location of the video camera for each of the apparatuses.

Order effects. To control for the effects of judging the gymnast's routines in a certain order, three different orders of the same routines were edited. Six judges saw order one, six judges saw order two and six judges saw order three.

Obtaining a Pool of Gymnastics Routines

Coach contacts for the clubs in the Montreal area were obtained at a regional competition at Gadbois Gymnastics Center. The coaches of the gymnasts were

telephoned, and videotaping dates were set up. Consent forms for the parents of the gymnasts were mailed to the coaches, and were filled out by the parents prior to the videotaping.

Videotaping. Each of the gymnasts was told to perform the routine as if she were performing at a competition. The gymnasts treated the video camera like it was the judges' panel and presented to the camera at the beginning and end of each routine. The gymnasts had 15 minutes to warm-up on the apparatus before they were videotaped performing their routine. This was approximately the amount of time allotted to warm-up in a competition setting. The gymnasts performed their routines on each of the apparatuses in the same order so that the researcher could keep record of the names. All the gymnasts performed on one apparatus before they all performed on the next apparatus.

After all the gymnasts were videotaped on all three apparatuses, they were videotaped for five seconds standing in a natural position with their arms by their side, two meters away from the video camera. This videotape position helped satisfy the question of whether the judge's aesthetic perception of the gymnast's body shape affected the score she awarded the gymnast.

Recording height and weight. After all the gymnasts performed on all three apparatuses, their height, weight and age were recorded and written in the table beside their name. The gymnasts from all clubs were weighed using a digital Sunbeam scale. They were measured for height by standing against the wall. A mark was made on the wall using a ruler. The height of the mark was measured to obtain the gymnast's

height in centimeters. The tables in Appendix G show the gymnasts' height, weight, age and BMI. For the purpose of anonymity the real names of the gymnasts are not used.

Routines used for the orders. Level D routines performed with no falls were used. Forty-two routines performed by 16 different gymnasts were used for the three sequential orders. The three videotaped orders contained the same 42 routines presented in different orders. Appendix H shows the routines numbers and the distribution of the routines across apparatus and BMI category. It also shows which routines were used as lead-ins, follow-ups and stimuli.

BMI ranking of the gymnasts. It was important to ensure that from the pool of routines there were an equal number performed by gymnasts from each quarter (low, medium-low, medium-high, high) of the Body Mass Index (BMI) of the gymnasts. The BMI is a ratio of height and weight which produces a proportionate value of height and weight indicating body shape for this study. BMI scores from the gymnasts' height and weight were calculated using the following equation:

$$\text{Body Mass Index} = \frac{\text{Mass}}{\text{Height}^2}$$

The gymnast's were then rank ordered by BMI. Four divisions of BMI were formed in order to randomly choose an equal number of routines from each division from the continuum of BMI for the videotape. The following table shows the range of body mass indices within each category.

Table 1- Appendix B

Body Mass Index Range within each Category

	Low (n=4)	Medium-low (n=4)	Medium-high (n=5)	High (n=5)
BMI	14.2-15.3	15.9-16.7	16.8-18.3	18.6-20.0

Organization and Numbering of the Routines

Three different orders of the same routines were used in the study. The routines were the same in each order, but appeared in different sequences to avoid the possible order effects. To do this, the videotapes of the unedited routines were viewed. A list was made with a number given to each routine. The position of the routine on the unedited videotape was recorded. All routines that had one or more falls were eliminated from the study. It was thought that a fall from the apparatus would bias the judges' scoring of the rest of the routine.

Random selection of routines for orders. Each routine was coded using a number, a shape for apparatus type, and a colour for BMI category. All routines were recorded on paper and placed into a hat. The first routine chosen from bars, from beam and from floor became the lead-ins, and the next routine chosen from bars, beam and floor were follow-ups. Lead-ins were added at the beginning of the formats to accustom the judges to the type of task that they were required to do and to prevent any confounding effects of scoring the routines that appeared first. The judges were told to

score these routines, but the scores were not used in the data analysis. The three follow-ups appeared at the end of the 36 routines, to prevent any effects on the scoring when the judge was aware of the termination of the one hour experiment (this procedure is similar to Ste. Marie & Lee, 1993). The rest of the routines became the stimuli. The same routines were chosen randomly from the hat three different times to form the three different orders. The lead-ins and follow-ups remained constant over the three orders, and only the order of the stimuli changed. Each order had three lead-ins, three follow-ups and 36 stimuli. The 36 routines were performed by 16 different gymnasts.

The order of the gymnasts for judge's perceptions. The order of the gymnasts standing in front of the camera for perception scoring was random, and was the same for all three of the orders of routines. The gymnast's names were put into a hat and drawn one at a time. For example, judge A saw order one of the routines and then saw the gymnasts in the anterior standing position in a certain order. Judge B saw order two of the routines and saw the gymnast in the anterior standing position in the same order as judge A.

Videotape Editing

Gymnasts' routines for the orders. All editing was done at the McGill Instructional Communications Center using a Panasonic editing system. After each routine, a 10 second blank was recorded in order to give the judge enough time to calculate the gymnast's score. At 10 seconds, a tone was recorded so that the judge could prepare for the next routine. One second later the next routine was displayed.

Gymnasts for judges' perceptions. Each gymnast was presented on the videotape from an anterior view standing upright for three seconds. After each presentation of the gymnasts, a five second blank was added to give the judge time to circle a whole number on the Likert scale. At five seconds, a tone was recorded so that the judge could prepare for the next gymnast. Five lead-ins were used at the beginning, and five follow-ups were used at the end of the sequence of gymnasts for the same reasons that they were added in the sequence of routines. Five fillers were also placed randomly among the stimuli. Fillers were added to vary the range of possible body shape perceptions scores, and to attempt to shift the judges' attention away from the purpose of the study. The scores for the lead-ins, follow-ups and fillers were not used in the data analysis.

The Pilot Study

Data Collection

After the editing of the first order was completed, data was collected from three beginner judges. Each judge was tested one day and then again three days later in order to test the judging consistency and the videotape reliability. The judges read and signed the consent form (Appendix I). The verbal instructions of phase one in Appendix E were read to the judge. The judges positioned themselves at a comfortable distance away from the television.

After the judge had finished evaluating the routines, the videotape was stopped and phase two of the verbal instructions (Appendix E) was read. When the judge had completed phase two of the task, the videotape was stopped. The judges were asked to fill out an experience profile. The judge was debriefed about the study verbally and any questions were answered. A written summary of the study was given to the judge for further information (Appendix J). Data from these judges were not included in the study.

Informal questioning

The judges were asked some informal questions about the task they had just completed. They were asked whether the visual background of some routines on the videotape was disturbing to the accuracy of their judging. They all responded that they were focused on the gymnasts, and the background had no effect on their ability to judge accurately. Two of the judges commented on how the routines were very clean (lacked technical errors). When asked about the task of rating their perception of the

gymnast's body shape, they responded that it was difficult because they had never done this task consciously. The judges were also asked whether they were given enough time to rate the gymnast's body shape for aesthetic appeal and whether they thought the angle of viewing was appropriate. One judge responded that she believed the idea was to give a quick impression of the gymnast's body shape and found that the task did just that. Another judge said that the angle of viewing the gymnast was similar to the angle that the judge sees as the gymnast presents before she performs and when the gymnast is finished.

When the gymnast presents, she raises both arms and with feet together stretches her body towards the judge. The gymnast presents at the beginning of her routine to indicate she is ready to perform and when she is finished to indicate the termination of her routine.

Results

Correlation of the scores for judging consistency. An intraclass correlation was performed on the gymnasts' mean performance scores and the judges' perception of aesthetic appeal of the gymnasts' body shape. The analysis produced an R of .99 for the performance scores and an R of .65 for the perception scores. The lower correlation for the perception scores was most likely due to the range of the possible scores that could be chosen by the judges for the gymnast's body shape perception score. The judges were required to choose a whole number from the Likert scale out of ten. If the judges circled a seven the first day, and a six the third day, then the correlation would not be high. However, the difference between these two scores is

minimal when speaking of the judge's perception of the gymnast's body shape. The judges found this task difficult as it is a task that they were not used to doing and this is probably the reason for the varying scores on the third day.

APPENDIX C

The Compulsory Routine Judging Sheet

Routine #1 Area for deductions Final score	Routine #2
Routine #3	Routine #4
Routine #5	Routine #6
Routine #7	Routine #8

APPENDIX D

Assessing the Gymnasts' Body Shape Sheet

Gymnast #1

least most
1 2 3 4 5 6 7 8 9 10

Gymnast #2

least most
1 2 3 4 5 6 7 8 9 10

Gymnast #3

least most
1 2 3 4 5 6 7 8 9 10

Gymnast #4

least most
1 2 3 4 5 6 7 8 9 10

Gymnast #5

least most
1 2 3 4 5 6 7 8 9 10

Gymnast =6

least most
1 2 3 4 5 6 7 8 9 10

Gymnast #7

least most
1 2 3 4 5 6 7 8 9 10

Gymnast #8

least most
1 2 3 4 5 6 7 8 9 10

Gymnast #9

least most
1 2 3 4 5 6 7 8 9 10

Gymnast #10

least most
1 2 3 4 5 6 7 8 9 10

APPENDIX E

Verbal Instructions Given to the Judges at the Time of Testing

Phase One

You will be presented with a series of routines on bars, beam and floor. A tone will be heard before every routine to cue you of the upcoming routine. During the routine you will need to deduct the appropriate marks for form errors and technical errors in the space provided on the judging sheets. After the routine is over, you will have 10 seconds to add up your deductions and subtract from a starting value of 10.0. If you find that the 10 seconds between the routines is not enough time, you can pause the videotape and finish the task. Do not rewind or pause the tape during any of the routines. You will judge exactly what you see.

Phase Two

Now you will view an anterior view of the gymnasts standing in one place. You will see each of them for three seconds. After you see each gymnast, you will ask yourself the question: At this level of gymnastics that you just saw, does this gymnast have a body shape that is aesthetically appealing for gymnastics? Then you must circle a whole number on the Likert scale provided that is appropriate for the gymnast. Ten is the most aesthetically appealing and one is the least aesthetically appealing.

APPENDIX F

Judges' Experience Profile Sheet

Name _____

1. How many years have you been involved in gymnastics as:

____ an athlete

____ a coach

____ a judge

2. Please list the gymnastics clubs you have been affiliated with over these years.

Club	Location (city)
_____	_____
_____	_____
_____	_____

3. Please explain what level you participate (ed) in as;

an athlete _____

a coach _____

a judge _____

APPENDIX G

Gymnasts' Age, Height, Weight and BMI

Figure 1: Gymnika/ Level D

Name	Age	Weight	Height	BMI category	BMI score
Amelie	10	59	130	med-low	15.9
Gabrielle	10	57	130	low	15.3
Kaitlin	9	58	136	low	14.2
Gail	10	69	137	med-high	16.8
Karyne	11	78	137	high	18.9
Shannon	11	86	144	high	18.9
Annie	11	78	148	med-low	16.2
Cindy	12	90	143	high	20.0
Nancy-Ann	11	90	150	med-high	18.2
Angel	11	66	136	med-low	16.2

Figure 2: IMCO/ Level D

Name	Age	Weight	Height	BMI category	BMI score
Roxanne	9	54	130.5	low	14.5
Marie	12	77	140.5	med-high	17.9
Lynn	10	67	135	med-low	16.7
Julie	13	95	152	high	18.7

Figure 3: Gymnacentre/ Level D

Name	Age	Weight	Height	BMI category	BMI score
Elaine	12	87	148	high	18.05
Paula	12	63	142	low	14.2
Joan	12	103	160	high	18.29
Emily	11	91	149	high	18.63

APPENDIX H

Distribution of Routines Across Apparatus and BMI Category

Bold = Stimuli

Italic = Lead-ins and Follow-ups

The numbers represent the number that the routine was given on the videotape

Low BMI

Bars	Beam	Floor
17	**14**	**12**
15	**10**	**3**
2	**2**	**2**
3	*8*	

Medium-low BMI

Bars	Beam	Floor
1	**1**	**1**
10	**6**	**7**
13	**9**	**10**
		11

Medium-high BMI

Bars	Beam	Floor
9	**3**	**4**
12	**11**	**9**
18	**15**	**13**
	13	

High BMI

Bars	Beam	Floor
5	**5**	**14**
22	**4**	**5**
8	**7**	**8**
6		*6*

APPENDIX I

Consent and Information Form for Judges

The research is being conducted by Sheri Valiquette. If for any reason you would like to contact me, the following information should provide you with the means to do so:

Sheri Valiquette
Graduate Student
McGill University
Department of Physical Education
Phone Numbers
in Ottawa (613) 726-1313
in Montreal (514) 288-7122

Purpose: The purpose of this study is to investigate differences in judgments of novice and experienced gymnastics judges. The results from this study will be ultimately used for the development of a more objective evaluation system in a sport that rely on judges to determine the outcome of a competition.

Demands: The experiment will require approximately one hour of your time. Your participation will involve viewing a series of gymnastics routines presented on video and performing different judging duties. If at any time you do not want to continue with the experiment, please make the experimenter aware of this and the experiment will be stopped. Any data collected will be destroyed and will not be used in this study.

Risks/Discomforts: There are no expected risks in this research. However, there may be slight discomfort by virtue of the fact that the experiment does require your full attention for the duration.

Anonymity and Confidentiality: The presentation of the data will be in a pooled format. That is, at no time will your individual scores be reported. In this way you have anonymity. In fact, we will not place your name on the data sheet and therefore will not be aware of any one judge's specific data results. This will ensure confidentiality of your results.

For your Information: At the conclusion of the experiment you will be provided with a handout describing in more detail the purpose of the experiment. The experimenter will be willing to answer any questions you may have concerning the experiment.

I, _____, have read the above information and agree to participate in this experiment.

Date: _____ Signature: _____

Experimenter: _____

APPENDIX J

Debriefing: A Bias for Body Shape

Thank you very much for having participated in my experiment. This sheet is designed to help you understand exactly what question I was investigating with this research.

Previous Research

Little research has been done on the biases involved with the gymnast's aesthetic look when doing gymnastics. However, at elite levels it is possible that this aesthetic component based on body type has a great influence on the already subjective judging process. Falls and Humphrey (1978) looked at body type differences between placers and nonplacers in a regional gymnastics meet. The physical measurements were done on 14 placers and 57 nonplacers in this AIAW meet as well as on 54 reference non-athletic females. The findings conclude that certain body types may be a factor in championship performance in women's gymnastics. "The precise explanation for improved gymnastics performance among women with this body type is not available from the research conducted thus far. It is possible that the neat ectomesomorphic body type conveys a more favorable visual image to gymnastics judges and somehow influences the scores awarded (Falls and Humphrey, p.43)."

As documented in Joan Ryan's book "Little girls in pretty boxes" from extensive interviews there is more evidence of body shape bias. The process for developing the U.S Olympic team was altered in order to allow Betty Okino, an athlete who was too injured to compete in either the U.S championships or the Olympic Trials, to have a chance to make the U.S Olympic Team. Betty's far from flawless performance at this second trial won over Kim Kelly's perfect performance and forced Kim from the Olympic Team (Ryan, 1995). Kim's coach recalls the selection committee's negative comments regarding the issue that would eventually doom Kim Kelly: body shape. Most commented about Kim's mature body of 5 feet 1 inch tall and 102 pounds in weight. "We want to walk onto the floor in Barcelona petite and strong," was one specific comment (Ryan, 1995).

The gymnastics judges have a difficult job of, for the most part, subjectively scoring the gymnasts' routines. Their perception of how a skill should look and how the gymnast should look doing the skill prevails. The stereotype of the gymnasts' body has long been a prejudice that everyone has accepted. Take, for example, two gymnasts: one is petit, 4'10" and weighs 90 pounds and the other is a little bigger, 5'1", and weighs 110 pounds. They can do the same skills, difficulty wise, and can execute them exactly the same. Who would get the better score? A bias for body

shape says that the smaller gymnast would because that gymnast is more pleasing to the eye.

Purpose of the Study

The purpose of this study is to determine whether judges at the club level of competition in gymnastics, give higher scores to those gymnasts that are shorter and weigh less. By having judges score compulsory routines shown on videotape and then later correlating those scores with the gymnasts' height and weight, the study will attempt to clarify whether gymnastic judges produce higher scores for those that are smaller and weigh less and lower scores for those gymnasts that are taller and weigh more. A second concern is whether the judges' perception of the gymnasts' body shape relates to the score that they give the gymnast for performance of her routine. Finally, the study will consider whether a bias for body shape differs for experienced judges (>10 years experience) in comparison with beginner judges (<5 years experience).

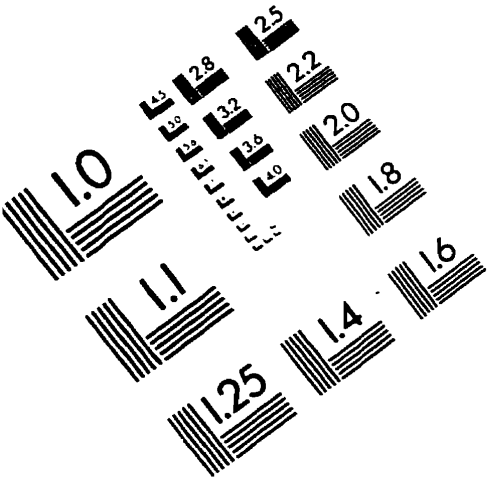
Every judge in this experiment performed the same task. You first viewed and scored compulsory routines of gymnasts with varying height and weight. Then you were asked to rate each gymnast's body. Using the scores and ratings, I can then use statistical procedures to relate the scores in order to examine the following hypotheses:

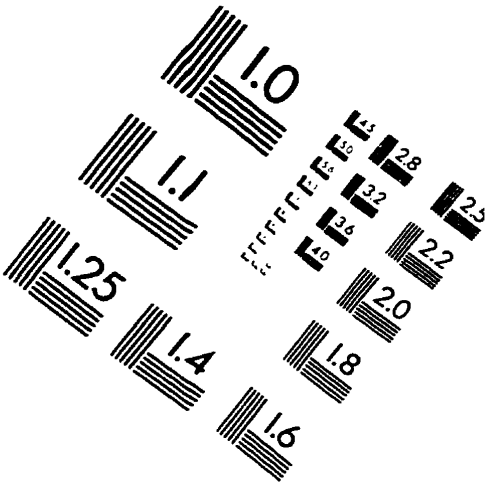
(1) The judges' perception of the aesthetic appeal of a gymnast's body shape will be higher for gymnasts who are short and weigh less than for those who are taller and weigh more.
(2) The judges' scores for gymnasts' performances will be higher for gymnasts who are short and weigh less than for those who are taller and weigh more.
(3) Experienced judges will rate gymnasts lower for aesthetic appeal than novice judges.
(4) There will be a significant difference between the gymnast's performance scores of experienced and novice judges.

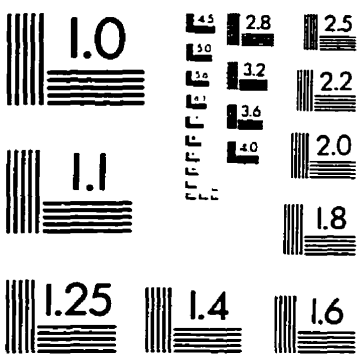
As mentioned in the information sheet, ultimately I hope to apply these results to the development of a judging system that eliminates biases that may hinder one's objectivity. If you have any questions, do not hesitate to contact the investigator listed below.

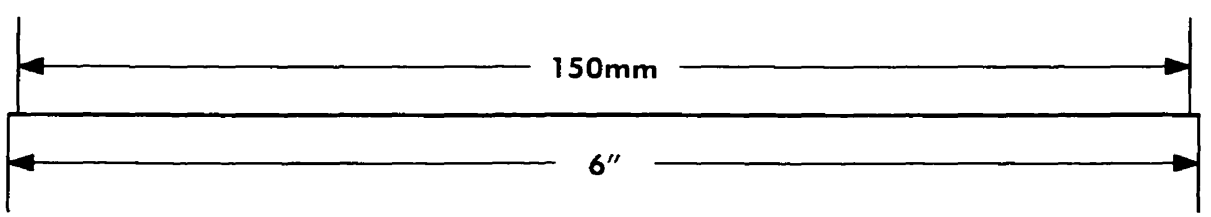
Sheri Valiquette
726-1313

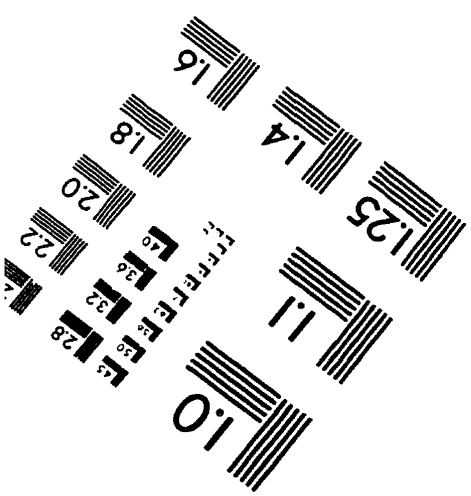
IMAGE EVALUATION
TEST TARGET (QA-3)

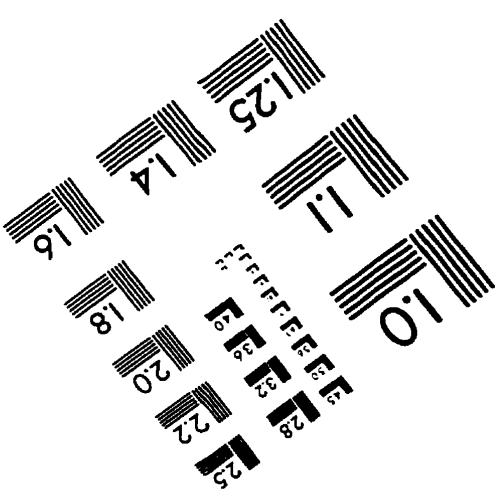












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved