

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.

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.

**ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]



Université d'Ottawa • University of Ottawa

**The Name Game: Is there a Reputation Bias in
Figure Skating Judging?**

Leanne Findlay

Thesis

**Submitted to the School of Graduate Studies in partial fulfilment of the
Master of Arts Degree in Human Kinetics**

School of Human Kinetics

University of Ottawa

2001



**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 référence

Our file Notre référence

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.

0-612-67817-2

Canada

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv

ABSTRACT v

CHAPTER I

 Revised Introduction 1

CHAPTER II

 Revised Literature Review 6

 Figure Skating Judging 7

 Social Cognitive Strategies for Perception and Evaluation 9

 The Effect of Expectations on the Performance Appraisal Process 13

 Stages of Performance Appraisal and Expectation Effects 17

 Encoding Stage of Performance Appraisal 17

 Evaluation Stage of Performance Appraisal 19

 Purpose and Direction of Research 20

 Relevance of the Research 21

CHAPTER III

 Presentation of Article to be Published 24

 Abstract 27

 Methods 32

 Participants 32

 Materials 33

 Procedure 36

 Dependent Measures 37

 Results 38

 Evaluation Phase of Judging 40

Encoding Phase of Judging 41

Discussion 41

Footnotes 47

Table 1 48

Table 2 49

Table 3 50

References 51

Author Note 55

CHAPTER IV

Elaborated Discussion 56

 Halo Error 58

 Factor of Personal Familiarity 60

 Limitations of the Research 62

 Future Research Directions 64

 Recommendations for the Applied Setting 65

 Conclusion 67

REFERENCES NOT INCLUDED IN THE ARTICLE 69

APPENDIX A

Contribution of Collaborators 72

APPENDIX B

Post-Judgement Questionnaire 74

APPENDIX C

Skate Canada Scoresheet (sample) 83

APPENDIX D

Instruction Set 85

APPENDIX E

List of Deductions for Errors in Novice Ladies Short Program 87

APPENDIX F

Consent Form and Information Sheet 89

APPENDIX G

Debriefing Sheet 92

APPENDIX H

Ethics Approval 95

Acknowledgements

I would like to gratefully acknowledge the assistance, direction, and support of Dr. Diane Ste-Marie, without whom the quality or content of this project would not have been possible. I would also like to thank Stephanie Brunet, Josianne Roma, Alison Cronin and Shannon Clark for their assistance and encouragement throughout the project. Finally, I would like to acknowledge the contribution of my committee members, Dr. Glen Kenny, School of Human Kinetics, and Dr. Pierre Mercier, School of Psychology, University of Ottawa. Apart from academia, I would like to express my gratitude for years of encouragement from my parents, Eldon and Kathleen Findlay, and to all of the other friends whom I have learned from and been supported by along the way. Finally, I would like to acknowledge the participants in the study, in particular Lauren Smith and Sylvie Parayre, who were instrumental in the coordination of the judges who gave their time and input to the study.

Abstract

Expectations for performance, such as an expectation set by within team order in gymnastics (Scheer & Ansorge, 1975; 1980), have been shown to create a non-performance bias in terms of sport performance evaluation. Plessner (1999) found such a bias to influence both the encoding and evaluation phases of gymnastics judgement. In the sport of figure skating, the encoding phase is considered the period during which a judge perceives elements and any errors that may occur. After the performance is complete, judges evaluate the merit of the performance by providing two marks that are combined to create an ordinal rank, which is considered the evaluation phase of judging. The current study investigated whether or not expectations created by the reputation, or name, of an athlete occurs in the sport of figure skating, and if so, in what phase of the process, encoding or evaluation. Fourteen female figure skaters' short programs were viewed by judges to whom the athletes were either known or unknown. Judges were asked to provide technical base marks, technical merit, and artistic impression marks (all indicative of the evaluation phase), as well as to identify the elements and associated deductions (the encoding phase variables). Post-experiment, the ordinal placements of the skaters were calculated and used to define whether biases do exist at a basic level. Indeed, the skaters' average ordinal placement was higher when they were known by the judge as compared to unknown ($t = 2.43, p < .05$). A more in-depth analysis revealed that skaters received higher technical merit and artistic impression marks, as well as technical base marks when known to the judges. No significant differences, however, were found for deductions allocated for errors or for the judges' percent correctness in perceiving the elements performed. Thus, the findings suggest that a reputation bias exists at the level of the evaluation phase, but not during the encoding phase, of figure skating judging.

Chapter I

Revised Introduction

Performance appraisal, in particular as it relates to sport, is a highly complex task.

Sporting events often rely heavily on subjective evaluation based on the perceptions of a judge, such as in the domains of gymnastics, diving, freestyle skiing, figure skating and so forth. Social cognitive processes can play an important role in such evaluation and hence should be included in studies of sport evaluation. In particular, how such processes may be influenced by expectations and prior knowledge of the performer (for example, his or her reputation) are of interest. Judges themselves would agree that there is a certain influence of reputation on the determination of an athlete's performance. For example, in the domain of figure skating, judges admit that the time limits imposed in judging are eased by prior knowledge such as reputation, but feel assured that any potential errors will be overridden by the other members of the judging panel (Ireland, 1996). If, however, a skater's reputation leads to a systematic bias, then the objectivity of the judging process has been compromised.

Expectations and prior knowledge are often cited by the media and the public as a biasing influence on performance appraisal. According to the popular press, suspect judging in ice dancing, for example, led to Canadian competitors Bourne and Kratz's fourth place, rather than a podium finish at the 1998 Winter Olympic Games (Deacon, 1998). In the world of competitive sport, subjectively judged events have continually been flagged as biased in terms of the objectivity of evaluation. This may be largely due to the fact that a number of biases have been shown to exist. For example, judges both in gymnastics (Ansorge & Scheer, 1988) and in figure skating (Seltzer & Glass, 1991) have been shown to favour performers representing their respective countries, i.e., reflecting an international bias. Although Ste-Marie (1996) tried to link this bias to an unconscious process on the part of the judge, no support was shown. Rather, it was

stated that international bias may be a technique used intentionally by judges to favour athletes from their respective country. Similarly, memory biases have been evidenced, whereby gymnastics judges' perceptual judgements are influenced by specific prior performances of a gymnastic element (Ste-Marie & Lee, 1991; Ste-Marie & Valiquette, 1996).

More specific to this research, expectation biases have been shown to play a role in the evaluation of sport performance. For example, Scheer and colleagues (Scheer, Ansorge & Howard, 1983; Scheer & Ansorge, 1980, 1975; Scheer, 1973) have repeatedly shown what has been termed a "within-team order effect". An expectation bias is created because judges expect coaches to select the weakest athlete to perform earliest in the rotation of one event within a competition, followed by the second weakest, and so forth, up to the best athletes competing last in the rotation. The result of this expectation bias is that judges score a routine competed at the end of the within-team order higher than if that same routine was evaluated early in the rotation. Expectations for performance have therefore been shown to be influential in the judgement process, but research has been limited in this area. It is the focus of this investigation to examine another possible expectation bias that could be influential on the judgement process, that of an athlete's reputation.

In addition to the limited scope within expectation bias research on sport performance, sport bias research has had little theoretical guidance. Cognitive processes which underlie or even guide the appraisal process have been overlooked as a basis for such biases. Thus, it is also the goal of this research to understand the potential expectation bias created by a reputation within a theoretical framework, in particular from a social cognitive perspective. This framework will provide a basis for discussion, thereby creating a grounding theme by which judges' cognitive

processes can be understood.

In terms of previous theoretical links, research on expectation bias has been associated with the cognitive appraisal process (Klaaren, Hodges & Wilson, 1994; Swann & Snyder, 1980) and to the psychological processes underlying such appraisal (Plessner, 1999). For example, within-team order effects have recently been accounted for by Plessner (1999) within the psychological concept of schema-based cognitive processing. Plessner argued that the order effect found in gymnastics judging was the result of a systematic bias generated by the use of social cognitive schemata. Expectation-congruent information was said to be the result of prior knowledge, leading to schema activation, which in turn led to biased perception and processing. Plessner found that expectations were influential in the judgement of gymnastic routines, and replicated Scheer and colleagues findings of a within-team order effect. It was also shown that a distinction could be made between the encoding and evaluation stages of performance appraisal, since the two were differentially affected by the expectation bias. Although Plessner's study provided insight into the judging process and expectation bias, to our knowledge it is the only study to date which makes a direct link between expectation bias in sport and the social cognition literature. Thus, one of the main purposes of this research is to study expectation biases in figure skating judging within the framework of social cognitive psychology.

In the following chapter, the interaction between sport performance appraisal and the effect of social cognitive processes will be examined. In order to link the theoretical framework to the task of figure skating judging, an outline of the critical components of the figure skating judging task will first be described. The components and effect of social cognition on sport performance appraisal will then be identified, in particular as they relate to judging. Connections

between cognitive knowledge and the objectivity, or lack thereof, of sport appraisal will also be discussed in terms of the various stages of performance appraisal.

Chapter II

Revised Literature Review

Figure Skating Judging

Overall, figure skating performance appraisal is a very complex and challenging cognitive task. The responsibility of performance appraisal rests with a panel of judges who observe performances and rate athletes based on the technical and aesthetic components of their program. It requires a great deal of attention and integration of information to determine scores in both the “objective” (i.e., technical merit) and more subjective (artistic) domains. Due to the often ambiguous nature of both the task and the scoring system, judgement relies heavily on the subjective perception and evaluation of athletic performance. The technical merit mark and artistic impression mark are combined to yield an overall ordinal score (or rank) amongst the group of competitors. In order to determine the technical score in the short program, a technical base mark is also created by the judge as a reflection of the technical difficulty of the elements in the program. Each mark is given on a continuous 6.0 scale, with the smallest difference being one tenth of a point. In the usual competition environment, there are two programs, a long program and a technical (short) program, the ordinals of which are combined to create the skaters’ overall competition placement. However, for the purpose of the current study, only the short program will be examined due to the greater objectivity of the scoring of this program over the long program.

For the technical program, skaters are given a list of elements that they must complete, and the quality and completion of these elements are scored by the judges. In order to generate technical merit marks, judges are provided with a list of deductions that correspond to specific errors that could occur within a given performance. Deductions are subtracted from the individual’s technical base mark, which is generated based on the difficulty of the program and by

the athlete's execution of these elements. As an illustration, a program with a triple-double combination jump would be considered to have a higher base mark than would a program with a double-double combination. The deductions given for a particular error, however, range in severity based on the degree of error. For example, if an athlete executes a jump with a two-foot landing, the deduction from the individual's base score is between 0.1 and 0.4 (Skate Canada, 2000, see Appendix D) and it is the discretion of the individual judge as to what amount is to be deducted. Judges become proficient at assigning deductions through training and subsequently are promoted throughout the judging ranks. Consequently, judges are given some freedom as to when to deduct, as well as to how much. For this reason, the deductions taken for certain errors were a dependent measure in the current study design.

The artistic impression score reflects the quality and style of the performer. Higher artistic marks are representative of greater interpretation of music, speed, flexibility, quality of edge and turns, etc. There are, however, no specific error deductions in this component nor are there specific requirements. Rather, the artistic impression mark is generated from a holistic perspective based on the judge's impression of the overall performance. This score will also be used as a dependent measure. In competition, a range for both the technical merit and artistic impression marks are set by the referee (on the judging panel) following the first skater's performance, and at the Novice competitive level these marks are often revealed immediately post-performance.

Judges are advised to take notes during the performance and are required to submit both the technical and artistic marks immediately after the individual athlete's performance, thus creating a time-pressured environment. At high-level competitions, judges receive feedback by

viewing all judges' marks after each skater; however, at the Novice level to be studied in the current investigation, not all competitions employ this open-marking policy. In addition, judges are provided with a scoresheet on which they record the elements being performed by the athlete, the deductions for these elements, and often associated notes. This immediate appraisal information is used to generate the final marks at the end of each performance.

As can be noted in this description of figure skating judging, there is a fair amount of complexity and high information processing demands involved with the judging task. In the next sections, a social cognitive information processing model will be presented that describes performance appraisal in a work setting. Although not perfectly suitable, this model was deemed appropriate for the development of the research question within the context of sport evaluation.

Social Cognitive Strategies for Perception and Evaluation

Making perceptual judgements can be a difficult task that involves the observation, integration, identification, and evaluation of a stimulus. Fiedler (1995, p. 136) argues that "in order to understand and explain social behaviour we need to take into account the hidden cognitive processes that mediate between external stimuli and overt behavioural responses". The depth of such a task is overwhelming if one considers the amount of information that an observer must take into account when recognizing and forming a judgement of a complex behaviour or event.

In order to minimize the time and complexity of evaluation, observers often use cognitive strategies, known as heuristics, to guide their search and to aid in the processing of the event. A heuristic can be defined as "[conservation] of cognitive energy in arriving at inferences and

impressions about people and events by using cognitive shortcuts” (Alcock, Carment & Sadava, 1994, p. 96). Heuristics can be based on such factors as representativeness, illusory correlation, availability, simulation, consensus with a previous experience, or a prior expectation of a behaviour or event. Tversky and Kahneman (1974) suggest that there are three principle heuristics by which the complexity of cognitive processing is reduced: *representativeness* of the input to a previous or consistent behaviour, the *availability* of a previous exposure to memory, and *adjustment or anchoring*, whereby perception is compared to a base exposure example. Their argument is that although these principles allow the cognitive task to be simplified, systematic errors or biases may occur as a result of dependency on the heuristic. For instance, Tversky and Kahneman refer to a bias that occurs as a result of the familiarity of a behaviour, whereby a salient past example would be used by the observer to be compared, or even alter a current event. The exhibited behaviour may be judged in relation to the anchor instead of a general standard, thus changing the judgement outcome. Although heuristics minimize the amount of transcription and integration of cognitive input and processing, biases may be an inherent part of this ease of processing.

Given that the performance appraisal process is a psychological assessment followed by a response, the influence of a cognitive shortcut is of great importance. In figure skating, a representation such as that which results from prior exposure or knowledge may serve as a heuristic to guide performance appraisal. As Tversky and Kahneman (1974) referred to, the heuristic may, however, elicit an altered response by the evaluator. The manner by which this process of evaluation occurs will be discussed in the proceeding section.

The stages of performance evaluation which integrate the use of mental heuristics with

social cognitive processing have been systematically outlined by DeNisi, Cafferty and Meglino (1984) in their model of performance appraisal. They suggested that performance appraisal is “the product of a set of cognitive operations which includes acquisition of information through the observation of performance, organization and storage of that information in memory, retrieval of information from memory, and its integration to form a judgement” (p. 361). This argument is confirmed by Feidler (1983), who suggested six similar stages of processing which are influenced by top-down, heuristic-based knowledge structures which create a response bias in judgement. Judgements and decisions, for example, are based on heuristics such as availability, representativeness, anchoring and simulation, but may come at the cost of “systematically biased judgements” (p. 148). DeNisi et al. (1983) argued that social perception and performance appraisal were strongly influenced by the attention that the rater placed on an active search for information. This attention process can determine both the type of information gathered by the rater (also termed encoding) as well as the categories of performance appraisal or the direct evaluation of the behaviour.

One of DeNisi et al.’s (1984) propositions that is important to this thesis was that an observer’s preconceived notions can influence many levels of the evaluative process involved in performance appraisal. More specifically, preconceived notions were argued to activate a schema (a heuristic or representation of a previous experience), which was then used to ease the recall of past performance and the integration of new information. These preconceived notions may be the result of past performances, for example, the consistency of the behaviour in previous instances may cause the individual to search for confirmation of this behaviour, thus eliminating the need for direct information search. Other research has supported the idea that preconceived notions on

the part of the performance evaluator may be responsible for setting up a heuristic that can be used to evaluate performance (Darley & Gross, 1983). In figure skating, it is conceivable to argue that preconceived notions based on prior knowledge or of information regarding a skater's reputation could exist.

Within expectation theory, expectancies are also described as being derived from prior knowledge or experience with performance evaluation (Feidler, 1995), and in this way are similar to DeNisi et al's preconceived notions. In fact, Miller and Turnbull (1986) argued that expectation biases may even persist in spite of contradictory evidence, such that behavioural confirmation in evaluation is achieved because of expectations. They argued that past experiences generate expectancies, which influences the processing of social behaviours. Thus, this literature review links together the concepts of preconceived notions and expectancy-driven biases.

Within the realm of social cognitive psychology, prior processing and knowledge are often linked to the concept of a schema, which according to Fiske and Taylor (1984, p. 84) is "a cognitive structure that contains the attributes of a concept and the relationships among those attributes". They argue that in order to simplify processing of social events, observers use a top-down process whereby prior concepts shape current perception. Top down processes are often used to explain ease of cognitive functioning, deriving from the Gestalt concept of holistic perception (Anderson, 1995). Top-down can be contrasted with bottom-up processing, which is the direct observation of data driven stimuli and the integration of these basic elements.

Similar to these top-down and bottom-up processes, Brooks (1987) distinguishes non-analytic and analytic processes. Non-analytic is akin to holistic perception, and analytic processing involves a more systematic and detailed search for input. Brooks argued that stimuli to

which the observer has previous exposure depend more on non-analytic processing since many elements of the stimuli are assumed to be unchanging; that is, the previously observed person or event has many constant elements which are perceived as one unit. This type of holistic processing allows the perceiver to process the information more quickly; however, it may lead to a condition which can be more systematically influenced by expectations or prior knowledge than that which would occur within non-analytic processing. In figure skating, this distinction would be similar to viewing each individual jump, spin, field move, etc.(bottom-up/analytical) versus viewing the performance first in a holistic fashion (top-down/non-analytically).

Linking this concept to figure skating further, the technical mark is based on individual elements performed by the skater; therefore, it could be argued to be more similar to the analytic style of processing. Conversely, given that the artistic mark is based on a holistic impression of the program, it is more like a non-analytic analysis of the performance. Perhaps, then, the influence of expectations may be differentially evident in the artistic and technical components of figure skating judging.

The Effect of Expectations on the Performance Appraisal Process

In the literature, it is shown how complete objectivity of identification or evaluation may be compromised by this ease of processing that comes along with the use of heuristics (Swann & Giuliano, 1987; Tversky & Kahneman, 1974). Two types of effects, assimilation and contrast, that can result from the use of the heuristic of using expectations to drive performance appraisal are described next.

Kingstrom and Mainstone (1985), found that a biased condition was created in a

workplace setting by the use of heuristics; ratings of work-place performance were affected as a function of the acquaintanceship of the rater and the ratee. It was determined that there was a significant effect of raters' personal acquaintance such that a higher degree of familiarity led to higher performance ratings for specific attributes as well as an overall evaluation. Thus, personal acquaintance between the rater and ratee influenced performance appraisal. A link between the use of mental heuristics as a guide in the search and evaluation of a ratee was also suggested. Although the work environment does not mirror sport, it is similar in that an authority figure must evaluate behaviour based on some form of measurement tool or perceptual standard. The subsequent rating is reflective of the ratee's behaviour in their performance environment.

Relatedly, Feidler (1996) suggested that positive behaviours have a greater influence on appraisal than do negative behaviours, such that a positive instance will be more likely to be recalled and thus alter performance appraisal than would a negative example. An illustration which translates to figure skating would be that a skater who had performed a difficult element once would be expected to be able to do that element in future, even if they had not been able to demonstrate it more than a few times. Thus, research demonstrates the effect which heuristics, especially positive examples of behaviour, can have on the evaluation process; prior cognitive knowledge may guide future reference because of expectations during evaluation.

Other research has also pointed to the effect of positive expectations and knowledge in influencing workplace performance appraisal (Foti & Hauenstein, 1993; Trafimow & Schneider, 1994). Swann and Snyder (1980) demonstrated that the perception of performance of a cognitive/skilled motor task was more strongly related to expectations rather than actual performance behaviour. In their research, randomly allocated "low"- and "high-ability" groups

were assigned to a card task. Despite the objectively superior performance of the “low-ability” group, the “high-ability” participants were perceived by the raters to have better performance results. This research thus identifies the significance of raters’ prior expectations, and how, even in the face of contradictory evidence, perceptual confirmation of behaviour can occur (see also Miller & Turnbull, 1986 for a review).

In order to explain the empirical evidence suggesting that expectations influence performance evaluation, two effects based on prior impressions have been formulated. The assimilation effect, which argues that subsequent judgements are made in the direction of prior performance, suggests that performance appraisal would be biased towards consistency with previous appraisals (see Murphy, Baltzer, Lockhart & Eisenman, 1985; Feidler, 1995). Smither, Reilly and Buda (1988) also suggested that prior behaviour was found to be correlated with an expectation. Thereafter, when evaluating that particular performer, the use of the expectation schema led to a tendency for the evaluator to believe a performance was consistent with the expectation. These results were also supported by those of Buda (1984, as cited in Smither et al., 1988), whereby raters given a positive example of teacher behaviour, i.e., positive prior knowledge, rated the subsequent ratee performance more favourably than did raters provided with negative prior performance.

By comparison, the contrast effect suggests that raters are more sensitive to the inconsistent elements of subsequent performances and are more influenced by these elements in the judgement process. For example, contrast effects would be indicated by a more harsh judgement when errors are made following a good performance. However, according to the assimilation hypothesis, these errors would be discounted, or perhaps even not detected, following

a good performance.

While Murphy et al. (1985) found support for both types of effects, other research has shown that the nature of the task demands can influence the type of effect evidenced. For example, in a study using videotaped lecture performances, Smither, Reilly and Buda (1988) found that assimilation and contrast effects were influential in the process of judging when raters were given both indirect (verbal) and direct (observational) knowledge, i.e., the creation of an expectation. They found, however, that contrast effects were eliminated when memory demands on the evaluator were increased and an element of time was introduced to the judgement process. Relatedly, Snyder and Stukas (1999) argued that perceiver's reliance on expectations became increasingly pronounced as cognitive load increased.

In the literature, Darley and Fazio (1980) found another feature of the task which was shown to be important in terms of the type of effect induced was the ambiguity of the stimulus. Assimilation effects were found to be particularly strong in affecting judgement when rater behaviour was ambiguous (Darley & Fazio, 1980). Thus, time pressure, increased cognitive load, and ambiguity have been shown to be three features of a task which lead to assimilation effects rather than contrast effects. The next section considers figure skating judging along these various dimensions.

First, as already discussed, figure skating judging has high cognitive demands. The judge is required to take note of the performance as it is being performed, as well as integrate all performed elements into one score upon the completion of the performance. In terms of time pressures during the performance, the judge has only milliseconds to perceive the athlete's behaviours as many of the elements are performed very quickly. In addition, a score has to be

generated immediately after the performance, before the next skater performs, and thus, there are also time pressures at this level. In terms of feature ambiguity, the nature of any errors can be difficult to discern in light of a varying angle of viewing, the degree of severity of an error, and the lack of clear definition of errors as they are made. In light of all the cognitive demands, time dependency, and task ambiguity, the assimilation effect would likely play more of a role in the judging of a sport such as figure skating. That is, the judge would rely on past behaviours to guide their evaluation of a current performance to be consistent with prior positive behaviours.

In sum, studies within social cognition demonstrate the role of expectations as a factor influencing cognitive strategies used by an evaluator during performance appraisal. In addition, an overview of the assimilation effect suggests that judgements will more likely be made in the direction of past performances. More information detailing the particular processing stages at which expectations can influence the performance appraisal process will be presented in the following section.

Stages of Performance Appraisal and Expectation Effects

Considering the complexity of social judgement, the evaluation process can be influenced by expectations at various stages. DeNisi et al.'s (1984) model, in fact, identifies six categorical steps of performance appraisal, however, for the current investigation these will be compiled into two main stages: encoding of information and the evaluation of performance. These stages were chosen because effects in the encoding and evaluation of performance information have been particularly identified in the literature. Encoding is represented by observation, organization, and storage of perceptual information and the evaluation phase of performance appraisal links to the

final three stages of the DeNisi et al. model, which are the retrieval and integration of information and the assignment of a formal evaluation. The next sections describe how these two components have been shown to be affected by expectations.

Encoding Stage of Performance Appraisal. In terms of the encoding process, it has been shown that prior knowledge may affect performance appraisal in a variety of settings such as workplace evaluation (Kozlowski & Ford, 1991), the teacher-student relationship (Rosenthal & Jacobson, 1968), personality studies linking traits associated with certain profiles (Swann & Giuliano, 1987) and sport studies (Plessner, 1999). Kozlowski and Ford (1991), for example, found that when the performance evaluator had a greater amount of prior knowledge about the performer, less active search was done on performance in order to form an evaluation, rather, “search appeared to be guided by confirmation of a general evaluation in memory (p. 298). von Collani, Kauer and Kauer (1991) also found that people use knowledge to guide an expectancy-driven search in the rating process. Once confirmatory evidence was obtained, the search was either discontinued or was misguided in the direction of the confirmed schema.

Another setting which has shown similar results concerning expectation influences is that of student performance appraisal by a teacher. In the teaching field, expectancy effects have been repeatedly documented (Anderson et al., 1991; Brophy & Good, 1970; Rosenthal & Jacobson, 1968). Brophy (1982, as cited in Anderson et al., 1991) went so far as to suggest that teacher expectations account for between five and ten percent of student achievement. Prior knowledge that a teacher has of the student, their grades, past successes/failures, etc., may serve to anchor any subsequent perception or judgement of the student’s behaviour. For instance, the teacher may be guided by an expectation during the pre-evaluative phase, especially if the prior knowledge was

thought to be reliable (Shavelson, Cadwell & Izu, 1977). This situation of prior knowledge is illustrated in figure skating whereby judges may be very familiar with an athlete's reputation, past successes/failures, and so forth. This information could conceivably set up an expectation bias on the part of the judge for the skating performance to be evaluated.

Indeed, expectation bias has already been identified in sport performance appraisal at the level of the encoding phase. Plessner (1999) reported that gymnastics elements were perceptually encoded incorrectly as a result of an expectation bias induced by within-team order (see also Scheer, 1973; Scheer & Ansorge, 1980, 1975). More specifically, if a gymnastic performance was scored later in the within-team order, error deductions (which are recorded on-line during performance, hence reflecting encoding) were not as severe as that given if the performance was evaluated earlier in the within-team order. Thus, Plessner's research reinforces the notion that expectations can affect performance appraisal at the level of encoding in sport performance evaluation.

Evaluation Stage of Performance Appraisal. In addition to altering the search pattern and encoding of perceptual information, expectations may influence the evaluation aspect of the observed behaviour that is being appraised. The effect on one stage does not preclude the effect on the other, nor does it guarantee that the other will also be affected (Snyder & Haugen, 1994). Both phases must therefore be investigated. Snyder and Haugen (1994) would suggest that expectations provide a framework for evaluation, and other supporting research has identified situations in which expectations may influence the evaluation process. For example, in a study on sport trait attributions during a cognitive task, Reeder (1997) found expectations of sporting ability had their greatest influence in the evaluation phase of judgement. In other words,

expectations had their effect during the time when observers had an opportunity to reflect on the overall performance. In figure skating, this period would be post-performance, in between the completion of the skater's routine and the awarding of a score. During this time, the judge reflects on the total performance by considering the information on the scoresheet recorded during the encoding phase and integrating this information with current knowledge to generate both technical and artistic merit marks.

The research done by Plessner (1999) also shows that this latter stage of the performance appraisal process can be affected in a sport setting. In order to assess the separate encoding and evaluation stages, Plessner (1999) first had gymnastic judges use a blue pen while they were recording the information as the gymnast was performing (the encoding stage). Immediately following the completion of a performance, the judges then switched to a green pen for post-performance notes/scoring. It was found that final marks were not directly calculated from the immediate evaluations. Judges' deviated from the marks that should have been given based on the encoding phase, thus demonstrating that performance evaluation was influenced by some non-performance factor, which in this case was argued to be the within-team order expectation bias. When judges were asked in a post-experimental interview why their marks deviated, they commented on the fact that their overall impression of the athlete caused minor changes to the marks. This relates to the top-down processing stated earlier, whereby the holistic interpretation of a behaviour may be influential in guiding or conceptualizing the formation of an evaluation. This type of processing, however, was skewed in the direction of prior knowledge, since the expectation was based on the knowledge that coaches place their strongest gymnasts at the end of a rotation. These combined findings of expectations leading to influences in performance

appraisal at both the encoding and evaluation stages of processing provide the necessary background related to the purpose of this research that will be described in the proceeding section of this chapter.

Purpose and Direction of Research

This research attempted to determine whether an expectation bias related to a positive reputation of a skater existed in figure skating. The research was framed within a social cognitive perspective of performance appraisal. More specifically, the purpose of the investigation was to determine whether or not a skater's "name" being associated with a positive reputation resulted in judges awarding higher marks over athletes whose names were not known. The two main questions addressed were: 1) does a reputation bias the performance appraisal process?, and if so, 2) during which cognitive phase(s) does this expectation bias have its influence? This bias, to our knowledge, has not yet been studied in any context, and is a phenomenon which is commonly commented upon in sport; however, it has not yet been proven with empirical evidence.

The hypotheses of the study were that a reputation bias would exist such that a judge's evaluation of a skater would be influenced by the athlete's name. This hypothesis would be supported if it was shown that athletes' final placements, as determined by their ordinal ranking, were better when their name was known versus when they were unknown. Further, both the encoding phase and the evaluation phase were expected to show a bias. For the evaluation phase, this would be revealed if the measures of technical merit, artistic impression, and technical base marks were greater for known athletes as compared to the athletes' marks when they were unknown. With respect to the encoding phase, if the variables of percent correctness in

identifying the elements and errors as well as the deductions allocated were lower when the athlete was known versus unknown, an expectation bias based on the name of the athlete would be shown. Systematic biases, like a reputation bias, hamper the objectivity of scoring in figure skating, and in turn jeopardize the semblance of impartiality in judging the sport. Considering that such a bias has not been included in sport research, an investigation is critical in order to gain a better understanding of the sport performance appraisal process.

Relevance of the Research

Given that the aim of this research was to study a potential bias in sport performance appraisal within a cognitive processing model, it can be said that both theoretical and applied ramifications of the research would be attained. On a theoretical level, past research in sport performance appraisal has neglected to integrate the psychological processes which accompany perception and evaluation. This research investigated these issues within a social cognitive framework, in particular with reference to a two-stage model of performance appraisal. Numerous researchers have investigated expectations as they apply to either perception or appraisal, however, research has inadequately probed the encoding and evaluation processes. The current investigation attempted to rectify this issue since both phases of performance appraisal process were studied in an applied context.

On a more applied and practical level, since proof for a reputation bias was expected, amendments to the process of performance appraisal would be suggested from the findings. The further investigation of judging bias, especially as it relates to subjective evaluation, will assist not only judges but also figure skating associations in terms of the development of the judging

process. For example, biases in the realm of ice dancing have been addressed due to scandalous uproar regarding blocked or unfair judging (Deacon, 1998). In much the same way, biases that exist in the freeskating discipline must be addressed. For example, loosely set deductions by which skaters are docked for errors may be problematic if judges are biased in their utilization and freedom within these rules. Only through knowledge can any corrections or improvements be made to equalize the playing field and eliminate non-performance advantages that some athletes may have over others.

A final point of mention is that subjectively judged sports often have a minuscule margin of difference in marks between athletes, indicating that even the smallest source of bias may have huge ramifications. For example, in the 1988 Winter Olympics, Brian Boitano defeated Brian Orser by only one judge and by one tenth of a point, the smallest possible margin of difference. If there was even an glimmer of bias in favour of Boitano, the elimination of such a bias may have meant the difference of a gold medal and millions of dollars in endorsements and contracts to Orser. The world of sport is filled with other such examples of small margins of defeat, and the recognition of biases in the process of judging is imperative if corrections for any biases are to be implemented.

Chapter III

Presentation of Article to be Published

Presentation of Article to be Published

The following paper titled “An Expectation-Induced Reputation Bias: Explaining its Effects on the Encoding and Evaluation Phases in Figure Skating Judging” will be submitted to a reputable journal within the field of Sport Psychology for publication. The abstract has also been submitted for an oral presentation at the SCAPPS conference 2001, to be held in October of this year.

Running Head: REPUTATION EXPECTATION BIAS IN FIGURE SKATING

**An Expectation-Induced Reputation Bias: Explaining its Effects on the Encoding
and Evaluation Phases in Figure Skating Judging**

Leanne C. Findlay

Diane M. Ste-Marie

University of Ottawa

Abstract

Expectations for performance, such as an expectation set by within team order in gymnastics (Scheer & Ansorge, 1975; 1980), have been shown to create a non-performance bias in terms of sport performance evaluation. Plessner (1999) found such a bias to influence both the encoding and evaluation phases of gymnastics judgement. In the sport of figure skating, the encoding phase is the period during which the judge perceives elements and associated errors; the evaluation consists of the creation of the technical merit and artistic impression marks. The current study investigated whether or not expectations created by the reputation, or name, of an athlete occurs in the sport of figure skating, and if so, in what phase of the process, encoding or evaluation. Fourteen female figure skaters' short programs were viewed by judges to whom the athletes were either known or unknown. 12 judges were asked to provide technical base marks, technical merit, and artistic impression marks, as well as to identify the elements and associated deductions, and post-experiment the overall ordinal placement of the skaters were calculated. Indeed, the skaters' average ordinal placement was higher when they were known by the judges as compared to unknown ($t = 2.43, p < .05$). A more in-depth analysis revealed that skaters received higher technical merit and artistic impression marks, as well as base marks when known to the judges. No significant differences, however, were found for deductions allocated for errors or for the judges' percent correctness in perceiving the elements performed. Thus, the findings suggest that a reputation bias exists at the level of the evaluation phase, but not during the encoding phase, of figure skating judging. Suggested implications of the findings suggest that education for judges, as well as changes to the figure skating judging process, are in order.

The task of evaluating performances in any realm is a complex and challenging task. The abundance of incoming perceptual stimuli must be taken into consideration as well as the integration of this information with previous knowledge, the evaluation criterion, and so forth. In order to simplify the task, performance appraisers often reduce the amount of processing needed to do the task by adapting certain strategies. Tversky and Kahneman (1974; Kahneman & Tversky, 1996) describe such strategies as heuristics, or representative tools that serve as short cuts for high processing demand situations. Although the use of heuristics may make complex judgement tasks more manageable and may increase the speed of processing, the problem exists that systematic biases may result. One heuristic that has been documented in the performance evaluation literature is the performance cue effect, whereby raters tend to make behaviour and evaluation ratings which rely on information based on prior behaviour and performance information (e.g., Baltes & Parker, 2000).

Using existing knowledge in this way is similar to setting up an expectation for the performance to be evaluated. Certainly, it is difficult to enter any evaluative social situation which is not prefaced in some way by expectations or assumptions about the various persons and characteristics involved in the situation. Such expectations have been shown, in both laboratory experiments as well as more naturalistic settings, to influence the performance evaluation process (for a review, see Miller & Turnbull, 1996).

Both positive and negative expectations can create a situation in which raters' subsequent interpretation of behaviours coincide with those particular expectancies. For example, it has been shown in workplace settings (Kingstrom & Mainstone, 1985) and in academic teaching evaluations (Murphy, Baltzer, Lockhart & Eisenman, 1985) that raters who were highly familiar

with the ratee tended to give more favourable overall ratings than when they were not familiar. Similarly, when supplied with a positive previous rating, and thus given a “positive expectation”, evaluators may use this expectation to generate future behaviour evaluation. For instance, Swann and Snyder (1980) demonstrated that the perception of performance of a cognitive motor task was more strongly related to expectations rather than actual performance behaviour. Raters were given an expectation that was randomly assigned, such that raters perceived one group to be of “low” ability and the other group of “high” ability. Despite the objectively superior performance of the low-ability group, the falsely-assigned high-ability participants were perceived by the raters to have better performance results. This research thus identifies the significance of raters’ prior expectations, and how, even in the face of contradictory evidence, perceptual confirmation of behaviour can occur.

To understand how such expectations function, it is useful to utilize an information processing framework that divides the perceptual appraisal process into different stages. Stages of performance evaluation which integrate the use of mental heuristics with social cognitive processing have been systematically outlined by DeNisi, Cafferty and Meglino (1984) in their model of performance appraisal. This model consists of six stages, which for the purposes of the current study will be collapsed into two levels. The first phase, termed the encoding phase, is represented by observation, organization, and storage of perceptual information. The evaluation phase of performance appraisal links to DeNisi et al.’s retrieval and integration of information and the assignment of a formal evaluation. Research has shown that influences due to expectation occur at both of these stages of the performance appraisal process (Plessner, 1999). First, expectations influenced the actual perception, or encoding, of the behaviour. In addition, the final

judgements provided after the evaluation phase were also shown to be influenced by expectations (see also Martell & Willis, 1993; Hauenstein, 1992; Darley & Fazio, 1980).

Although these expectation effects have been shown quite consistently in the domain of management (Kingstrom & Mainstone, 1991) and teaching (Rosenthal & Jacobson, 1968), another context in which they may exist, but has received much less attention, is the realm of sport performance. Certainly, biases in general have been demonstrated in sport, such as international biases (Whissell, Lyons, Wilkinson & Whissell, 1993; Seltzer & Glass, 1991, Ansorge & Scheer, 1988), memory influenced biases (Ste-Marie & Lee, 1991; Ste-Marie & Valiquette, 1996; Ste-Marie, Valiquette & Taylor, in press) and within-team order bias (Scheer, 1973; Ansorge, Scheer, Laub & Howard, 1978; Scheer & Ansorge, 1975, 1980; Plessner, 1999). The latter bias, in fact, has been explained as an expectation bias, wherein expectations held by gymnastic judges were argued to lead to a systematic bias to score routines performed in the end portion of a within-team order higher than those in the first portion of the order. The expectation was said to be induced by a common strategy used by coaches, that of placing their strongest athletes later in the within-team order of the rotation. Thus, judges expect that gymnasts who perform later in the rotation to be better than the earlier athletes.

Although several biases have been shown to exist, one that has yet to be investigated is the potential impact of a bias as a result of the reputation, or the "name", of an athlete. Indeed, Swift (1998) has suggested that a judge's performance appraisal of a skating program may be influenced merely by the reputation of the athlete. In fact, Otta, Vanzolinida Silva Leme, de Penha Pinheiro Lima & Rocha Sampaio (1983) did find a name-related bias in the evaluation of a text composition. Specifically, they found that a text accredited to Freud was rated higher than if the

same text was reported to be authored by Skinner or if no author was given. The fact that more favourable judgements occurred solely because of the name of the author would suggest that expectations in figure skating could also result from the reputation, i.e., the names, of the skaters. If a skater is “known” to have a positive reputation, this may construct an expectation for a good performance, as compared to a situation in which the skater is unknown.

Thus, the purpose of the current study was to determine whether a reputation bias based on the name of an athlete exists in figure skating. At this general level, we investigated whether overall rankings (ordinals) of figure skaters in a competition could differ as a function of whether skaters’ performances were evaluated by a judge that knew their name, versus those that did not. Ordinal rankings were used at this general level because they are used in competition to determine skaters’ final placements. Our hypothesis was that the ordinal ranking of figure skaters would be better when the athlete was known to the judges as compared to when they were unknown.

In addition, we wanted to identify the particular stage(s), encoding or evaluation, during which such a bias might influence judges’ ability to objectively appraise a performance. To do this, we identified specific measures for each of the two phases. First, the encoding phase was defined as the period during which the judge observed the program and took note of the particular elements performed and the error deductions associated with each element. These two measures were thus used to determine whether the encoding phase was influenced by expectations. The evaluation phase was assumed to occur upon the completion of the skater’s program, during the time frame when the judge integrated all of the performance information into a final judgement in the form of a technical base mark, a technical merit mark and an artistic impression score. Thus, these three marks were used as the evaluation phase measures.

Our hypotheses were that the evaluation phase of performance appraisal, as measured by the technical and artistic marks as well as the technical base marks assigned to the skater, would influence whether the athlete was known or unknown to the judges. Specifically, we proposed that better scores would be attained when the athlete was known versus unknown. In terms of the encoding measures of percent correctness and the deductions for performance errors, we hypothesized that judges' would make fewer observational errors (indicated by their percent correctness) and would accredit more deductions to unknown skaters than known skaters. These hypotheses were made despite the fact that some management research has shown that encoding may not be effected by expectations (Murphy, Balzer, Lockhart & Eisenman, 1985), because Plessner (1999) reported that in gymnastics, the within-team order expectation bias to influence both the evaluation and encoding phases of performance appraisal variables. Considering the similarities in terms of the social environment of sport, it can be argued that in the figure skating setting, judges will be similarly influenced at both phases.

Methods

Participants

The participants were six qualified judges from the Eastern Ontario section of Skate Canada¹ and six from the Quebec section (total n = 12). The selection of 12 judges was based on the fact that the skaters' marks when they were known and unknown were averaged across judges. All judges volunteered to be in the study and provided written consent before being tested. The average age was 45 years, and 5 of the judges were male, 7 were female. For the purpose of the study, participants were required to be qualified by Skate Canada to judge at least the Novice Competitive level (the level of the skaters presented); however, 10 of 12 were qualified to judge at

the Senior Competitive level and 6 were qualified for National or International-level competitions . The judges had an average of 15 years experience, with a low of 7 years and a high of over 30 years judging.

Materials

To set up the experimental stimuli, videotapes of live competitions were obtained of the 2000 Eastern Ontario Sectional event from the Eastern Ontario Sectional office and of the 2000 Québec sub-Sectional competitions from the Québec videotaping company “ES & Nat, Enregistre”. The short program (which is 110 to 130 seconds in duration and is composed of eight specific elements) was used due to the more objective standards with regards to required elements and deductions. In order to select skaters for inclusion on the research tapes, three figure skating judges from each section were asked to rate a list of all skaters² that had been collected on videotape in terms of their reputation in their respective section (total number of skaters was 81).

A questionnaire was used to discern whether or not the skaters had a name (or positive reputation). This questionnaire asked the question, “Do you think that the skater has made a positive name for herself within the Section?”. To classify the skater as known in the skating community, she needed to obtain a three or four out of a maximum score of four. Thus, the main purpose of the questionnaire was to identify skaters who had made a name for themselves in the skating community and who would be known to most, if not all, members of that community. Further, only skaters that were identified by at least two out of three judges as “known” were used for the experimental videotapes.

In total, seven skaters were identified as “known” by Ontario judges and seven skaters were also identified as “known” by the Quebec judges. The number of skaters was selected to be

14 due to fact that the judgement task had to limited to a time which was reasonable to demand of the judges. Of significance is that to the Ontario judges, half of the skaters were thus known skaters (the Ontario athletes), whereas those same athletes were unknown to the Quebec judges. Conversely, to the Quebec judges, the seven Quebec athletes were known and the Ontario skaters unknown; this distinction was checked during the experimental protocol. Thus, when all of these programs were randomly ordered and edited onto one tape, it created an equal number of known and unknown skaters being judged.

In addition to these fourteen skaters, one additional skater was presented on the experimental videotape. One skater's (who was unknown to both groups) program was used to set the "start-score" relative to which all other programs would be marked and was therefore not included in the analysis. This first program also allowed the judges to become familiar with the research protocol. The technical and artistic "start scores" were created prior to the experiment by having two judges evaluate the first skater's short program. These judges had the opportunity to rewind and review the performance as many times as desired. The technical and artistic marks from these two judges were then averaged to create the starting technical and artistic marks for that first program viewed by the judges.

The post-experiment questionnaire consisted of three questions. The first question was the same as that used to determine whether skaters were known or unknown in the skating community. That is, judges were asked, "Do you think that the skater has made a positive name for herself within the Section?". The questionnaire also asked the judge to: 2) Please rank your personal familiarity with the skaters below (a Likert rating of 4 was regular interaction), and 3) Before today, how many times do you remember having judged this particular skater over their

skating career?. (The effect of being familiar with the program and not the skater directly was assumed to be negligible).³ Since the first question dealt directly with the reputation bias, it was the only question used in order to confirm that the skaters were “correctly” identified by the in- and out-of-province judges.

A second video tape was created for a presentation of the same fifteen skaters to the judges to assist them in completing the post-experiment questionnaire after they had finished judging the 15 short programs. More specifically, a 10 second clip of each skater was created whereby the skater was shown presenting herself to the judges either before or after her program. In this way, the clip was used to stimulate the recall of the judge so that he/she would know what skater they were being asked to reflect upon when responding to the questionnaire.

The Skate Canada score sheet typically used by judges in actual competitions was provided to the participants. To eliminate any possible regional bias that may exist in the judgement process (see Ansorge & Scheer, 1988 regarding international bias), two formats of the scoresheet were created with the only difference being the names of the skaters to be evaluated. More specifically, for the Ontario judges, the list of skaters presented on the scoresheet contained the actual names of the seven Ontario skaters and fictitious, “English-sounding” names for the seven Québec skaters. Similarly, for the Québec judges, the Québec skaters were identified with their actual names whereas fictitious “French sounding” names were used for the Ontario skaters. In this way, judges were led to believe that all skaters were from the same general region, thus creating the known/unknown manipulation without varying provincial or regional representation. Judges were also allowed to consult their own personal competition deduction sheet for scoring the athletes, a practice which is customary in competition.

Procedure

Judges were recruited via telephone or electronic mail and asked to participate in the research. Participants were given the option to undergo testing either in their home or in a neutral location such as an arena or community center; all judges chose to be tested in their home. Prior to the procedure, judges were told that the research purpose was to study cognitive processing in sport evaluation, without other specifics given about our interest in reputation biases. Participants were, however, fully debriefed at the completion of the experiment concerning the research question of expectation biases and were given the opportunity to be excluded from the study. None of the participants elected to do so.

Participants were asked to sit an appropriate distance from the television in order to adequately see the screen. All screens used were at least 19", although most judges viewed a 27" screen. Once seated, an overview of the research procedure was given, as well as the variables that the judge would be asked to provide after each performance. These variables were: technical base mark (a reflection of the technical merit mark barring no errors), technical merit mark, artistic impression mark, identification of the elements performed and perceived deduction values for elements. Each judge was instructed that they should not be concerned about "blocking-in"⁴ athletes with marks. Because we wanted the judges to deal with the programs on a more individual level, we asked them to not consider the dimension of placement during the judging process.

Following this explanation, judges watched the first skater's program and were provided with the start scores of that program. This ensured that all judges were evaluating the programs from the same basic standard. At this time, they were also given the opportunity to ask any

questions. After this first program, judges then evaluated the next 14 programs on the videotape and provided information on all variables on the scoresheet. As in actual competition, judges were allowed a short period after each skater's performance in order to complete the judging process. As well, in keeping with regular practice, judges were not permitted to rewind or review the tapes. After marking all 14 programs, the researcher examined the judge's scoresheet and conferred with the judge, if necessary, to ensure that the judge's personal shorthand could be properly read for later scoring of the data.

Participants also completed the post-experiment questionnaire concerning each skater's reputation and the judge's personal acquaintance with the skaters. The stimulated recall video was used to remind the participants of each individual skater before they completed the questionnaire for that particular skater. The purpose of the questionnaire in this stage of the research was to ensure that judges were in fact unfamiliar with out-of-province athletes and that in-province athletes were considered known. If the judge's response did not coincide with the proper classification, their marks for that particular athlete were discarded from the analyses.

Dependent Measures

Three measures were used to assess the evaluation phase of judgement: the technical base mark, the technical merit mark, and the artistic impression mark. The technical base mark is generated after the skating performance and is a "starting technical value" for the program based on the difficulty as well as quality of the elements. The technical merit mark is a reflection of the execution and cleanliness of the elements in the program, and is generated from the base technical mark, taking into consideration the appropriate deductions for errors (Skate Canada, 2000). The artistic impression mark, which can be considered the more subjective of the two marks, reflects

the aesthetic components of the program, such as the speed, power, and beauty of the program.

In addition to the evaluation phase measures, two marks for identifying the encoding phase were analysed: 1) percent correctness in identifying that the elements were perfect or contained an error, and 2) the value for the error deduction given when errors were detected. To generate the standard for the percent correctness measure, three independent raters viewed all 14 programs and completed an element-by-element analysis to determine whether an error was associated with the specific skating elements or if the performance was free of errors (i.e., perfect). Only if the inter-rater reliability was found as 100 percent was the element included in the analysis for the dependent measure of percent correctness. Thus, not every element was used for each skater's program. Percent correctness was calculated by examining the congruity between these independent raters' judgements and that which the judge gave. For example, if the judge did not indicate a deduction was necessary for an associated error, yet one was noted by all three raters, the judge's perception was deemed as incorrect. As well, if a judge indicated that an error was present, however, the raters did not view a necessary deduction, that element was also noted as incorrect. Conversely, the detection of an error that was also identified by the three raters, as well as the judgement of a perfect element as in fact perfect, were scored as correct.

Another important measure was the ordinal rankings for the skaters that was calculated by the researcher after the judge had scored all programs. Similar to the actual skating environment, ordinal ranking is determined by summing the technical and artistic marks in order to create a ranking, with the technical score breaking any ties in the short program (Skate Canada, 2000).

Results

To analyse the data, marks for all skaters were averaged across judges, such that each

skater received an average value for when they were known to judges versus when they were unknown to judges. Each skater's marks were only used if her identification as known or unknown by the judge in the post-experiment questionnaire was in line with the manipulation of the protocol, i.e., the Ontario skaters were identified by Ontario judges as known and Québec skaters as unknown and vice versa for the Québec judges. One skater was completely eliminated from analysis due to the fact that she was unknown to all judges but two. Also, for two of the skaters, four judges' scores were used to create the average known marks. This occurred due to the fact these skaters were not identified by two judges that they were known, despite being from the same region. Likewise, one out-of-province skater was identified by a judge as known; therefore, five judges' marks were used to create the average unknown marks. Also, due to the multiple analysis of scores, a Bonferroni adjustment was applied and alpha was accepted at the $p < .01$ level for the evaluation and encoding variables.

In order to determine if a reputation bias existed, a paired t-test comparing final ordinals for when the athletes were known to the judge versus when they were unknown was conducted. This analysis revealed a significant relationship such that the average placing of the athlete was lower (i.e., better) when the skater was known rather than unknown $t(12) = 4.11, p < .05$. (See Table 1 for grand total means and standard deviations.) These results show that when the skater was known to the judge, they were, on average, more favourably placed by one and a half placement points, thus revealing a reputation bias. This ordinal placement, however, does not identify the phase during which the name bias influenced the judging process. Therefore, we undertook a more detailed analysis of the particular components which create this value by examining the variables associated with the evaluation and encoding phases.

Evaluation Phase of Judging

The evaluation phase was examined by submitting the technical merit and artistic impression marks to separate paired-samples t-tests. For technical merit, results indicated that skaters were given a significantly higher mark when they were known to the judge than when they were unknown to the judge, $t(12) = 3.30, p < .01$. (To see the evaluation and encoding variable results subdivided by province, see Tables 2 and 3). The artistic impression mark was found to be marginally significantly higher when the skater was known to the judge versus when the skater was unknown, $t(12) = 2.31, p = .04$. Thus, the skaters were given on average 0.17 higher for technical merit and 0.11 higher for artistic impression when they were known to the judge as compared to when they were not known, for a total of 0.28 difference.

Although not typically shown in an actual competition, the placement of the technical merit and artistic impression marks were also analysed. Both the ordinals created by the technical merit marks and the artistic impression marks were deemed to be significantly different for known versus unknown. Technical merit placement was on average one placing lower (i.e., better) when the skater was known to the judge versus when they were unknown to the judge ($t(12) = -3.24, p < .01$). Similarly, artistic impression placement was also significantly better ($t(12) = -2.64, p < .01$) when the skater was known to the judge than when they were unknown.

Another evaluation phase variable which was identified as a potential area in which bias might occur was the technical base mark given to a skater. It was found that when the skater was known to the judge, they were given a significantly higher base mark than when the skater was unknown ($t(12) = 3.25, p < .05$). Thus, known skaters were given a base mark which was 0.1 higher than if they were unknown to the judge.

These indicators of the evaluative phase of judgement showed that skaters' marks and placement can be affected by whether or not they were evaluated by judges who were aware of their reputation, such that if they were known by the judge they received higher technical, artistic, and base marks than if they were unknown. Since these marks are constructed post-program, they indicate that during the evaluation phase of the judging process, marks were influenced by the factor of being known.

Encoding Phase of Judging

In order to test the idea that judging bias may have also occurred in the encoding phase of the judgement task, percent correctness for the judgement of the elements was analysed. A paired-samples t-test showed no differences between groups (see Table 3), indicating that the judges were not significantly different in terms of perceiving errors for known versus unknown athletes.

Data were also analysed in terms of the deductions allocated for errors. These deductions are indicated directly on the scoresheet and can range in severity, typically from 0.1 to 0.4. Thus, the total deductions given for an entire program were calculated for the skaters' performances, and were then analysed using paired t-tests to compare the deductions when the skater was known versus unknown. Although greater deductions were shown for unknown skaters as compared to known skaters, this difference was not found to be significantly significant. Overall, both the findings for deductions and percent correctness would suggest that the encoding phase of the performance appraisal process was not affected by a skater's reputation to be known in the skating world.

Discussion

The two main purposes of this study were to determine if an expectation bias could result

from the name of an athlete, and secondly to explore the particular stage of evaluation in which such a reputation bias might have an effect. The general finding of this research was that a reputation bias does exist in figure skating judging. The ordinal placement, which is used to determine the competition ranking, was better when the athletes were known to the judge than when they were unknown. Thus, in conjunction with similar research (Plessner, 1999; Binning, Zaba & Whattam, 1986), the notion that evaluators can base their score on non-performance factors was shown here. This coincides with our original hypothesis that an expectation bias due to a positive reputation, as identified by the name of an athlete, would exist.

Given that the reputation bias was shown at the general level of ordinal placement, investigation into the two phases of performance appraisal allowed for us to discern how such a bias was cognitively mediated. The phases of interest here were: the evaluation phase, when the judge renders a decision as to the score of the athlete; and the encoding phase, during which time the judge perceives information from the behaviour to be assessed and writes this information on a scoresheet. Previous sport-related research has shown that an expectation bias can occur at both levels of evaluation (Plessner, 1999), and thus we hypothesized this to also occur here. Performance expectations, however, were found to only influence the evaluation phase, in particular the technical merit mark, of performance appraisal and not the encoding phase. The evaluation phase measures show that marks for the known/unknown scenario amounted to almost 0.3 in the overall score, which could have a large impact on the athlete's final placement.

In order to understand the differences in effect on the technical, and not artistic, scores, it is critical to recognize the varying types of processing which is occurring. According to Smith (1989), stimuli can be perceived and categorized either by an analytic process, one of deliberate

cognitive effort to identify constituent properties, or by a holistic approach, which is a more non-strategic, impressionistic dimension of processing. Expectations come into play when prior events serve as a guide in dealing with a current event. The main problem with holistic perception is that the whole is assumed to be consistent with the previous instance, such that differences are not as easy to identify (Brooks, 1987). In the context of our study, holistic processing is represented by the artistic impression score, and the evaluation of the constituent elements, i.e. the technical merit mark, is analytic due to the deliberate cognitive tasks of perceiving and recognizing elements and errors in performance. The findings of the study, that the technical marks are significantly affected by reputation bias, would suggest that the analytic processes are influenced by the expectation created by the name of the athlete.

Why is the same skating performance being given higher marks when it is judged by judges that know the name of the athlete versus when she is unknown? It is our contention that the “name” of the skater was associated with a positive reputation that led the judge to expect the known athletes to perform solid and aesthetically superior skills. This expectation thus sets up a “reputation bias” and higher technical and artistic marks are given in order to maintain consistency with expectations derived from previous successes (Klaaren, Hodges & Wilson, 1994; Swann & Snyder, 1980). More specifically, similar to Martell and Willis’ (1993), arguments for explaining a performance expectation effect at the level of evaluation suggest that the decision criterion of the judges was systematically influenced by the reputation bias. In other words, judges were more liberal when giving the base mark, technical and artistic marks to those skaters with whom a positive reputation expectation existed.

Turning to the encoding phase of the judgement process, our measures were related to the

on-line observation of the execution of elements and the deductions noted for errors. These two variables were not found to show a significantly different score dependent on whether the athlete was known or unknown. A trend, however, was found in terms of the deductions awarded such that the program received fewer deductions when evaluated by judges that knew the athlete as compared to when the skater was unknown. Perhaps a larger sample size would have made this effect more powerful. Despite this, equivocal results have been shown in previous studies concerning the impact of expectations on the encoding phase of performance appraisal. Baltes and Parker (2000), for example, reported that the encoding phase was not affected by pre-performance expectations. In contrast, Swann and Guiliano (1987) argued that the encoding phase was influenced by a confirmatory search strategy whereby the observer searched for information which coincided with a previously identified profile. Kozlowski and Ford (1991) also found that raters who already had an impression of the worker to be rated engaged in less of a search for information. These findings in others settings suggest that future research using other variables for the encoding phase, such as visual search patterns, might be worthwhile.

Another important point of mention is that Plessner (1999) found that expectation biases did exist at the level of encoding in gymnastics judging but that the effect was specific to the complexity of the elements being judged. More specifically, more difficult elements showed the bias yet easier elements did not. This would suggest that future research with respect to the complexity of the skating elements evaluated in the judgement task would also be useful in determining whether a reputation bias exists at the encoding phase. For instance, the difficulty in assessing a triple-triple combination jump, which consists of two, three rotation jumps in less than two seconds, is much greater than judging a double-rotation jump. Similar to Plessner's work, the

increased difficulty might lead to great errors in judgement as a result of expectations..

Given that a reputation-based expectation bias was found, potential strategies to reduce such a bias are important to consider. In terms of suggestions at the structural level of the judging process, the creation of fixed base marks, based on element inclusion, in addition to distinct deductions (versus ranges) may be in order. Currently, there is a range in severity in terms of the amount of deduction given for any type of error, and some freedom is allowed such that judges can award a different deduction score for the same error in performance by two different skaters. The potential for bias could be reduced by having a finite deduction, whereby the judge would be accountable to a specific error. Another possibility for reducing bias may be to eliminate the use of names, despite the fact that judges may be familiar enough that they recognize the appearance of a skater. Some bias, however, might be reduced by requiring them to have specific recollection of the athlete rather than a cue (name) to trigger expectations.

Although we did use a more real-life task in this research, several limitations can still be noted, some of which provide impetus for future research. First, the judgment task was done on video and thus the experimental setting loses several factors prevalent in a live competition environment. For example, the speed and quality of elements may have been more difficult to identify than in the natural setting; therefore, judges may have been forced to rely on any existing previous knowledge of speed, for example, in order to facilitate scoring. In addition, some elements and deductions were made increasingly ambiguous due to the two-dimensional nature as well as the difference in perspective on television.

Also to note is that previous research has shown factors such as panel-type feedback (Wanderer, 1987), the psychological characteristics of the judge (Scheer, Ansoerge &

Howard, 1983), and the specific form of the expectation, such as verbal accounts versus previous visual exposure to the athletes (Darley & Gross, 1983), affect how and if an expectation bias will occur. None of these factors, however, were considered in the current investigation, and therefore future research on the reputation bias should take these factors into consideration. For example, Plessner (1999) found that the within-team order effect occurred during the perceptual phase only when judges evaluated gymnasts in a panel environment, meaning that feedback issues may have been interacting with the effect. Wanderer (1987) also found that the artistic impression score in figure skating was particularly influenced by the presence of feedback. In the current study, judges were asked to evaluate in an individual situation and not as a panel, thus eliminating this feedback issue which may potentially influence the expectation bias researched.

If one considers the small margin of error and the closeness of marks in figure skating, the implications of any type of bias are apparent. For example, a “new” athlete, be that a person who is unknown in one area or who is new to the scene, may be negatively affected based merely on the non-performance factor of a lack of reputation. The implications are considerable for not only that particular athlete but also for the sport at large if an impartial means of evaluation is taking place. As was found in this study, the effect of having a positive reputation, or a name in the sport, is significant such that a reputation can cause inflated (or in the case of an unknown athlete, deflated) marks. This reputation-based expectation bias is of serious consequence for athletes, and as such should be identified and closely monitored in order to minimize the effect on performance outcomes in the sport of figure skating.

Footnotes

1. Skate Canada is the governing body for amateur figure skating in Canada. In order to reach Novice qualification, judges must attend clinics and seminars, as well as gain a significant amount of experience before being promoted by Skate Canada to this level.
2. The list for the Eastern Ontario skaters consisted of 16 names, whereas for the larger Québec Section the list consisted of 55 skaters.
3. It was assumed that the judges would have been exposed to that particular program very few times, if at all. This is due to the fact that the elements change annually; therefore, most skaters have a new program, performed to different music, in each competitive season. Since the number of competitions per season is minimal, the chance for any one judge to have evaluated a program many times before this exposure was very unlikely. The reason this was monitored is that we wanted to minimize any potential “mere-exposure” effects that could arise for programs seen repeatedly by the same judge (for a review of these effects, see Bornstein, 1989).
4. Blocking-in is a term used when judges attempt to categorize performances and assign marks based on order rather than performance. If blocking-in occurs, the judge would give an athlete a score with the purpose of placing them below or above another skater and not merely based on performance factors.

Table 1

Mean Scores for Skater's Ordinal Ranks Grand Mean and Shown by Province of Origin

		Average Ordinal	SD
Ontario	Known	6.78	0.58
	Unknown	6.71	0.51
Quebec	Known	6.94	0.70
	Unknown	6.49	0.57
GRAND MEAN	Known	6.85	0.67
	Unknown	6.6	0.58

Table 2

Mean Scores on Evaluation Variables Shown by Province of Origin and Grand Means

		Technical	<u>SD</u>	Technical	<u>SD</u>	Artistic	<u>SD</u>
		Base Mark		Merit		Impression	
Ontario	Known	3.72	0.23	3.08	0.37	3.70	0.23
	Unknown	3.67	0.22	3.01	0.36	3.68	0.20
Quebec	Known	3.81	0.23	3.14	0.54	3.83	0.16
	Unknown	3.66	0.19	2.87	0.50	3.61	0.10
GRAND	Known	3.76	0.25	3.11	0.48	3.76	0.22
MEAN	Unknown	3.66	0.22	2.94	0.44	3.65	0.17

Table 3

Mean Scores for Encoding Variables Shown by Province of Origin

		Percent Correct	<u>SD</u>	Deductions	<u>SD</u>
Ontario	Known	88.98	9.05	0.68	0.23
	Unknown	91.86	7.50	0.66	0.14
Quebec	Known	86.22	7.37	0.68	0.34
	Unknown	84.67	7.15	0.78	0.34
OVERALL	Known	87.73	8.77	0.68	0.27
	Unknown	88.54	8.50	0.72	0.27

References

- Ansorge, C. J., & Scheer, J. K. (1988). International bias detected in judging gymnastics competition at the 1984 Olympic Games. Research Quarterly, *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.
- Baltes, B. B., & Parker, C. P. (2000). Understanding and removing the effects of performance cues on behavioural ratings. Journal of business and psychology, *15*, 229-246.
- Binning, J. F., Zaba, A. J., & Whattam, J.C. (1986). Explaining the biasing effects of performance cues in terms of cognitive categorization. Academy of Management Journal, *29*, 521-535.
- Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising from the social interaction sequence. American Psychologist, *35*, 867-881.
- Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labelling effects. Journal of Personality and Social Psychology, *44*, 20-33.
- DeNisi, A. S., Cafferty, T. P., & Meglino, B. M. (1984). A cognitive view of the performance appraisal modal and research propositions. Organizational Behavior and Human Performance, *33*, 360-396.
- Hauenstein, N. M. A. (1992). An information-processing approach to leniency in performance judgements. Journal of Applied Psychology, *77*, 485-493.
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. Psychological Review, *103*, 582-591.
- Kingstrom, P.O., & Mainstone, L. E. (1985). An investigation of the rater-ratee

acquaintance and rater bias. Academy of Management Journal, 28, 641-653.

Klaaren, K. J., Hodges, S. D., & Wilson, T. D. (1994). The role of affective expectations in subjective experience and decision-making. Social Cognition, 12, 77-101.

Kozlowski, S. W. J., & Ford, J. K. (1991). Rater information processes: Tracing the effects of prior knowledge, performance level, search constraint and memory demand. Organizational Behaviour and Human Decision Processes, 49, 282-301.

Martell, R. F., & Willis, C. E. (1993). Effects of observers' performance expectations on behaviour ratings of work groups: Memory or response bias? Organizational Behaviour and Human Decision Processes, 56, 99-109.

Miller, D. T., & Turnbull, W. (1986). Expectancies and interpersonal processes. Annual Review of Psychology, 37, 233-256.

Murphy, K. R., Balzer, W. K., Lockhart, M. C., & Eisenman, E. J. (1985). Effects of previous performance on evaluations of present performance. Journal of Applied Psychology, 70, 72-84.

Plessner, H. (1999). Expectation biases in gymnastics judging. Journal of Sport and Exercise Psychology, 21, 131-144.

Otta, E., Vanzolinida Silva Leme, M. A., da Penha Pinheiro Lima, M., Rocha Sampaio, S. M. (1983). Profecias Auto-Realizadoras EmSala de Aula: Expectativas de estudanted de psicologia como determinanates não-intencionais de desempenho. Psicologia, 9, 27-42.

Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the classroom. New York: Holt, Rinehart and Winston, Inc.

Scheer, J. K. (1973). Effect of placement in the order of competition on scores of

Nebraska high school students. Research Quarterly, 44, 79-85.

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

Seltzer, R., & Glass, W. (1991). International politics and judging in Olympic skating events. Journal of Sport Behaviour, 14, 189-200.

Skate Canada. (2000). Skate Canada Rule Book. Ottawa: Author.

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

Ste-Marie, D. M., Valiquette, S., & Taylor, G. (in press). Memory biases occur under different prior processing conditions. Research Quarterly for Exercise and Sport.

Ste-Marie, D. M., & Valiquette, S. M. (1996). Enduring memory-induced biases in gymnastic judging. Journal of Experimental Psychology, 22, 1498-1502.

Swann, W. B. Jr., & Giuliano, T. (1987). Confirmatory search strategies in social interaction: How, when, why, and with what consequences. Journal of Social and Clinical Psychology, 5, 511-524.

Swann, W. B. Jr., & Snyder, M. (1980). On translating beliefs into action: Theories of ability and their application in an instructional setting. Journal of Personality and Social Psychology, 38, 879-888.

Swift, E. M. (1998, February 16). Blind Justice. Sports Illustrated, pp. 52-54, 57, 60.

Tversky, A., & Kahneman, D. (1974). Judgement and uncertainty: Heuristics and biases. Science, *185*, 1124-1131.

Wanderer, J. J. (1987). Social factors in judges' rankings of competitors in figure skating championships. Journal of Sport Behaviour, *10*, S93-S102

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

Author Note

At the time of completion of the article, Leanne Findlay was a graduate student in the School of Human Kinetics at the University of Ottawa. Dr. Diane Ste-Marie is an Associate Professor in the School of Human Kinetics, University of Ottawa. The first author completed this experiment in partial fulfilment of her Masters of Arts degree, and wishes to acknowledge support received from the SSHRC grant of the second author. We would like to extend thanks to Stephanie Brunet and Josianne Roma for their assistance with translation, as well as to Shannon Clark and Alison Cronin for their support throughout the research process. Finally, we would like to acknowledge the figure skating judges who participated in the research.

Address for correspondence: Diane M. Ste-Marie, 125 University Avenue, School of Human Kinetics, University of Ottawa, ON, K1N 6N5, CANADA, dstmarie@uottawa.ca

Chapter IV

Elaborated Discussion

Elaborated Discussion

The next section contains results and discussion of items and points of interest that were not included in the article. This chapter will include a correlational analysis of the skaters' technical and artistic marks, and between the technical base mark and technical score, as well as a discussion of those results in terms of the halo effect. A point of interest regarding the relatively low degree of accuracy in perceiving the elements of performance will also be mentioned. Following this, a discussion regarding holistic and analytic processing will draw together the literature reviewed in the second chapter, as well as the application of the findings. Finally, limitations and applications of the research for future applications, both in terms of research and the applied setting, will be elaborated upon.

Halo Error

Research has shown that halo error can strongly affect performance appraisal, such that individuals are rated as consistently good or bad performers based on a general impression that the rater holds for that performer (see Saul, Downey & Lahey, 1980 for an early review). Bernadin (1977; in Baltes & Parker, 2000) suggested that halo occurs because raters place a given ratee similarly on different dimensions. In figure skating, a halo has been shown to exist such that the judge believed the skater to show high quality performance, either technically or artistically, and the second mark was thus influenced as well (Weekley & Gier, 1989). For example, a skater could perform all of the technical elements well and receive higher artistic marks than warranted based on their technical abilities instead of aesthetic components. Although the two marks can be argued to overlap somewhat, for instance, a fall on a technical element can detract from the flow of the program and thus reduce the artistic mark, their independence should still be evident.

An interesting effect was discovered from the study; a significant correlation occurred between the technical and artistic marks for the skaters' programs. In fact, the known skaters' technical and artistic marks were more strongly related ($r = .79, p < .01$) than were the marks when the skaters were unknown to the judges ($r = .58, p < .05$). This would suggest that the judges may have been more strongly influenced by the halo effect, that global knowledge of the skater influenced specific ratings, in the condition in which they knew the performer. Weekley and Gier (1989) also found figure skating judges' technical and artistic marks to be highly related for the short program, reasoning that accuracy may be comprised by halo error. In this study, the high correlation between the two marks indicates that a high mark for one score can strongly predict a high score on the other.

Similar to the relationship found between technical and artistic marks, a strong significant correlation was also discovered between the technical score given and the technical base mark ($r = .90$, $p < .001$). Although this relationship also makes intuitive sense considering that the technical score is derived from the base mark, due to unpredictable errors in the program, it would be expected that the marks would deviate to a higher degree. If in fact the base mark reflected purely the technical content of the program, the merit marks would likely be more varying once the skaters performed and executed different errors. Such was not the case, indicating that perhaps the judges created the technical base mark or the technical merit marks more on knowledge of the performer than on the actual content. Considering that significant differences were not found between the deductions of known and unknown skaters, the judges may have given a higher base to the known skaters in order for their technical score to remain higher than unknown skaters, despite deductions.

Turning now to the encoding variable of percent correctness in identifying the element performed, when the analysis was conducted, an interesting fact was noted. The judges were found to be correct in identifying errors less than 90 percent of the time, meaning that errors in perceiving the elements were common. In fact, it was found that judges were 87.7 percent correct ($SD = 8.7$) when they knew the athletes, and 88.5 percent correct ($SD = 8.5$) when they did not. This would suggest that some errors are in fact common in the judgement process, and this finding alone would question the consistency in judging as well as the accountability for the results. Future directions for research might include investigating ways by which the degree of accuracy in observing elements and errors could be improved upon.

Factor of Personal Familiarity

Of interest is that when the factor of personal familiarity of the judge with the skater was removed, differences between the artistic marks when the athlete was known or unknown was even further reduced. This is contrary to the predictions made in Chapter 2, that the athletes' artistic scores would be greater when the skater was known versus unknown to the judges. This finding would suggest that judges were perceiving the artistry of the program itself rather than of the individual skater, and were thus not affected by reputation. If they were scoring the program, they may or may not like that piece of music, choices in choreography, etc., whereas for analytic, or technical, knowledge, they may use expectations to guide their decisions about things that would be expected to be more stable such as technical ability. If the athlete was technically able to execute an element one time, the technique will appear the same in every instance. (It is much more difficult to change the technical execution of an element than it is to change the presentation of the element, for example, by changing the entry into the element, the landing position, the artistry between elements, etc.) In this way, the judge may assume consistency in technical ability, and thus use their expectations in determining the technical score, whereas the holistic artistic impression score may vary more and depend less on their expectations.

Limitations of the Research

Several limitations of the study were previously mentioned in the article; however, they will be explained in more depth in the following section. One major concern of the participants was that the evaluation of skaters was done by watching a videotape, which is different from the actual competition environment. In particular, the participants commented that they had difficulty

identifying the speed of the skaters (a component of the artistic score), and also experienced some difficulty evaluating particular elements, for example, the circular step sequence. The problem that may have arose is that if the skater was known, the judge may have used past experiences of the skater to note speed, whereas unknown skaters would not have this advantage. This could inflate the score of the known skaters and disadvantage the unknown athletes, merely due to the videotape procedure.

Secondly, the construction of videotapes, as opposed to judging live performances, may have altered the judging process. For instance, the stress level of the judge must be considered if one is to explain cognition, in particular if cognitive processes are examined. The time-pressure environment of figure skating judging was not equally stressful with a live event, and thus the influence of the reputation bias may occur differently when they judge must render a score immediately and with feedback from having the score posted. Also due to the design of the study, the impact of “boxing-in” or tying performances was reduced. If, for example, a judge would base the decision in a tie situation to place one athlete over another based on the skaters’ reputation, the effect would have been eliminated in this study. By using videotapes and contriving the research study design, some of the important factors that occur in a live setting were compromised and therefore future research could investigate the possibility of using live performances to look at the research question.

The third limitation of the study was that judges reported that they had difficulty interpreting the post-judgement questionnaire in which they were asked to differentiate between how well they knew the skaters and how many times they had judged them. In fact, some judges commented that they were familiar with the skater; however, they did not rank the skater as

known to themselves on the questionnaire. This would suggest that some ambiguity may have existed in the questions posed in the questionnaire, and for future research it is advised that these types of questions be specific or quantified in some way.

In terms of the generalizability of the findings, several limitations must also be noted. The first area of concern is that of feedback. Plessner (1999) found that the within-team order effect occurred during the perceptual phase only when judges evaluated gymnasts in a panel environment, meaning that there was some amount of feedback or self-presentational concern interacting with the effect. Feedback was also shown to be critical in Wanderer's (1987) research on observed conformity effects in figure skating judging. Wanderer found that the artistic impression score in figure skating was particularly influenced by the presence of feedback, which would suggest that holistic perception may be more prone to influence by the evaluation of others. The impact of feedback may therefore be an important, yet neglected factor in the research process due to the laboratory-based nature of the research. In the current study, judges were asked to evaluate in an individual situation and not as a panel, thus eliminating this feedback issue which may potentially influence the expectation bias researched.

A second issue which limits the generalizability of the research is that of the psychological characteristics of the judges themselves. Scheer, Ansorge and Howard (1983) found that the within-team order expectation bias was related to the traits of dominance, autonomy and locus of control such that high feelings of all of these variables led to a decreased expectation effect. Within the current research project, however, these personality traits were not measured, and may be of interest for future investigations of expectation biases in figure skating.

Also of importance in expectation research is that the nature of the task may play a role in

terms of the effect that reputation bias may have on performance appraisal. Plessner (1999) noted that different gymnastic elements were affected by perceptual or evaluation biases in different ways. It was discovered that on the faster, more difficult apparatuses, judges' marks illustrated a within-team order placement effect, whereas the less difficult tasks were not so influenced. This would suggest that the complexity of the task may be a factor in the existence of an expectation bias. This might also occur in figure skating, wherein different elements may be differently affected by the judges' prior knowledge or expectations.

Finally, expectancy confirmation may also be limited by the type of prior knowledge the observer has of the performer. Raters may have prior knowledge based on two main sources: 1) personal experience in judging or watching the performer, or 2) vicarious knowledge through discussion or contact with other judges who have had experience(s) with that particular performer. Darley & Gross (1983) found that an expectation was insufficient to produce the behaviour-confirmation bias without prior personal performance knowledge. That is, bias was not shown when judges were given a verbal description of the participants' prior performance; the effect occurred only when the evaluator had actually viewed previous behaviour. This suggests that physical encounters and/or direct observations may be necessary to produce the effect.

Although the personal experience a judge has with a particular skater was noted, the main factor used to determine a skating reputation was the skater's notoriety to judges. This evaluation was based on previous successes, experience, potential, and so forth; however, this may have been skewed by the particular judges asked to complete this task. The generalizability of the research may be constrained by such limitations; however, investigation is essential if sport performance appraisal and the cognitive processes associated are to be truly recognized and understood.

Future Research Directions

Most research involving an expectation set derives these expectations by informing the raters of certain aspects of the ratee just prior to the behaviour to be assessed. In addition, the expectation set is often given only once. A strength of the current study was that the expectation set was built over a long period of time and was perhaps more salient and personalized. The benefit derived from the applied nature of this research is thus suggested for future research directions, making such research more ecologically valid.

The current study demonstrated one type of expectation bias in sport; however, future research should focus on other expectation biases that may exist. Also, although the encoding and behavioural rating phases were investigated, it is important to further understand the phases of cognitive appraisal in other subjectively judged sports. Considering that most of the expectation bias research in sport has concentrated on the sports of figure skating (Wanderer, 1987) and gymnastics (Scheer & Ansorge, 1975, 1978, 1980; Plessner, 1999), other sports are in need of investigation pertaining to specific areas of the judgement process that may be influenced by expectations.

The fact that only trends were seen for the encoding phase may suggest that a larger sample size, or another means of measure, are necessary in order to determine if encoding biases exist as a result of reputation. For example, the encoding phase could be investigated by measuring visual search strategies, such as the time spent viewing the pertinent elements of the target behaviour. A more detailed investigation of search strategies used by judges during encoding could identify whether the on-line processing of perceptual information was affected. (The search process has previously been shown to be influenced by expectation biases in

workplace management studies (see Kozlowski & Ford, 1991; Swann & Guiliano, 1987)).

Another line of research that is of interest to the researchers is that of the differences in processing which exist between expert and novice judges. Previous literature has shown that expert raters are significantly better than novice judges at anticipating upcoming gymnastics elements and that those elements that were anticipated were judged more correctly (Ste-Marie, 1999). Thus, the expectation bias that has been shown due to the reputation of an athlete could be investigated in that context in order to determine whether such an effect can be seen for either group.

Finally, a concern of the participants was that they did not have the opportunity to see skaters' warm-ups, which they suggested allows them to preview or prepare for upcoming elements and potential base marks and deductions. Although it was insisted that they were "not pre-judging", but merely using the warm-up information to prepare for upcoming programs, this experience may trigger previous exposure and memories, thus exacerbating an expectation bias. In fact, Ste-Marie and Valiquette (1996; see also Ste-Marie & Lee, 1991) found that gymnastic judges were biased by performances that were viewed prior to the actual competition, indicating the memory-influenced biases do indeed exist.

Recommendations for the Applied Setting

The main applications of the research in terms of real-life settings include educating judges about the influences that expectations can have on performance appraisal and implementing strategies to avoid any subsequent biases. Since a reputation bias was shown to exist, more elaborate criteria to determine the technical, artistic and technical base marks may be necessary. For example, the marks assigned for artistic impression are particularly ambiguously

awarded based on the broad qualifications of speed, quality and aesthetic performance, and may need more defined marks and deductions in order to reduce the impact of bias. Technical merit marks and base marks are also loosely defined, and could be more regulated if there were specific base marks given for particular elements in the program. As an illustration, a technical program which includes a certain combination jump, such as a triple-double versus a double-double, might be assigned a specific “start value”, similar to the judging practices in the sport of gymnastics.

The use of video-replay, which has recently been introduced at the World level of international figure skating competition, may also be of assistance in scoring programs. If the judge is made more accountable to their score, and the technology is available for use, they will perhaps be less reliant on expectations and knowledge sets when the behaviour to be judged is ambiguous.

Judge training may also increase the accountability of the sport performance appraisal process. In fact, rater training strategies, which identify the existence of biases to the rater, have been shown to be an effective means to reduce bias in student ratings of instructors and workplace performance appraisal (Pulakos, 1984; Bernardin, 1978 in Baltes & Parker, 2000). For instance, Baltes and Parker (2000) found that halo error training and structured recall memory tasks of previous performances were both effective means for reducing the performance cue effect (i.e., the effect of using prior knowledge when making a behavioural decision). Since figure skating judges already partake in a plethora of training in order to obtain their qualification, halo error training could be included at this time. Merely by making judges aware of the existence of the halo error might lead them to take greater caution to segregate the technical worth of the program from the aesthetic component when forming a judgement. In terms of structured recall memory,

figure skating judges could be asked to review the list of competitors prior to judging the event, and asked to recall previous experiences, both positive and negative, with specific skating events. If the judge were to recall a wide range of performances, they might remember the skaters' successes as well as failures, and would thus be more inclined to carefully view the current performance.

Another means by which Skate Canada could reduce expectation biases based on previous experiences and knowledge is by reducing, or eliminating, opportunities in which the skater is known. For example, by having judges from an outside region evaluate competitions, the reputation of the athletes would be unknown to the judges. The practicality of this suggestion, however, will obviously be of different magnitude in the case of Senior level competitors (who compete in the international sphere of competition) versus the Novice level studied here since Novice competition is more regional. At the elite Senior level, skaters are known on a world-wide basis, and finding judges who are unfamiliar with such skaters as Elvis Stoyko, for example, would be impossible and thus the reputation expectation set could not be eliminate.

Conclusion

The influence of a reputation bias on the outcome in figure skating has considerable ramifications if one considers the small degree of difference between the marks of the athletes in competition. Any bias that may result from non-performance factors can seriously impede the occurrence of just and fair sport judgement, and as such, expectation biases should be studied for both the theoretical determinants of perceptual processes as well as the changes necessary for the applied setting to be improved. In this study, a reputation bias was shown to exist in figure

skating, and the impact of bias on two major phases were identified: the encoding and the evaluation phases of judgement. In terms of the evaluation phase of appraisal, the technical merit score, the artistic impression score, and the base marks were all shown to be higher when the athlete was known to the judge. Although the encoding phase also demonstrated a trend such that deductions were less for known athletes, significance was not reached. In terms of the theoretical implications, the findings would suggest that expectations interfere with performance appraisal at the level of decision making and not necessarily perceptual encoding. On an applied front, any advantage which is not related to performance, but rather is the result of a bias, is of concern, and as such should be considered a handicap in the sport that must be rectified.

References Not Included in the Article

- Alcock, J. E., Carment, D. W. & Sadava, S. W. (1994). A textbook of social psychology. (3rd Ed.). Scarborough, Ontario: Prentice-Hall Canada.
- Anderson, J. R. (1995). Cognitive psychology and its implications. New York, New York: W. H. Freeman and Company.
- Brooks, L. R. (1987). Decentralized control of categorization: The role of prior processing episodes. In U. Neisser (Ed.) Concepts and conceptual development: Ecological and intellectual factors in categorization. New York: Cambridge University Press.
- Brophy, J. E., & Good, T. L. (1970). Teachers' communication of differential expectations for children's classroom performance. Journal of Educational Psychology, 61, 365-374.
- Canadian Figure Skating Association. (1995). Judging . . . An inside look. [Brochure]. Ottawa, ON: Author.
- Deacon, J. (1998, March 2). Dirty dancing. Maclean's, pp. 40, 42.
- Feidler, K. (1995). Processing social information for judgements and decisions. In Hewstone, M., Stroebe, W., & Stephenson, G. M. Introduction to social psychology (2nd ed.). Cambridge, Massachusetts: Blackwell Publishers Inc.
- Fiske, S. T., & Taylor, S. E. (1984). Social cognition. Reading, Massachusetts: Addison-Wesley Publishing Company.
- Foti, R. J., & Hauenstein, N. M. A. (1993). Processing demands and the effects of prior impressions on subsequent judgement: Clarifying the assimilation/contrast debate. Organizational Behavior and Human Decision Processes, 56, 167-189.
- Ireland, J. (1996, March 19). The jury's still out on judging. Ottawa Citizen, p. C12.

Pulakos, E.D. (1984). A comparison of rater training programs: Error training and accuracy rating. Journal of Applied Psychology, 69, 581-588.

Reeder, G. D. (1997). Dispositional inferences of ability: Content and process. Journal of Experimental Social Psychology, 33, 171-189.

Scheer, J. K., & Ansorge, C. J. (1980). Expectations in judging. International Gymnast Technical Supplement, 1, TS1-2.

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

Shavelson, R. J., Caldwell, J., & Izu, T. (1977). Teachers' sensitivity to the reliability of information in making pedagogical decisions. American Educational Research Journal, 14, 83-97.

Smith, J. D. (1989). Analytic and holistic processes in categorization. In B. E. Sherp & S. Ballesteros (Eds.), Object perception: Structure and process (pp. 297-323). Hillsdale, NJ: Erlbaum.

Smither, J. W., Reilly, R. R., & Buda, R. (1988). Effect of prior performance information on ratings of present performance: Contrast versus assimilation revisited. Journal of Applied Psychology, 73, 487-496.

Snyder, M., & Haugen, J. A. (1994). Why does behavioural confirmation occur? A functional perspective on the role of the perceiver. Journal of Experimental Social Psychology, 30, 218-246.

Snyder, M., & Stukas, A. A. (1999). Interpersonal processes: The interplay of cognitive, motivational, and behavioural activities in social interaction. Annual Review of Psychology, 50,

273-303.

Trafimow, D., & Schneider, D. J. (1994). The effects of behavioral, situational, and person information on different attribution judgements. Journal of Experimental Social Psychology, 30,

351-369.

von Collani, G., Kauer, A., & Kauer, S. (1991). Use and activation of social knowledge in commonsense attribution, evidence for a schema-based attribution model. Psychological Reports, 69, 728-730.

Weekley, J. A. & Gier, J. A. (1989). Ceilings in the reliability and validity of performance ratings: The case of expert raters. Academy of Management Journal, 32, 213-222.

Appendix A

Contribution of Collaborators

Contribution of Collaborators

Dr. Diane Ste-Marie suggested articles to review for the purpose of selecting a research area for my Master's thesis, and after reading on several topics, I chose to investigate expectation biases in judging. From personal experiences in the sport of figure skating, I chose to study whether or not the name of an athlete, in other words, the reputation of an athlete, was in fact influential on judgement as has been speculated upon by athletes, coaches, and the media. With Dr. Ste-Marie and the lab group's assistance, a methodology by which we could investigate this question was determined. Specific hypotheses were then formulated from various literature bases in terms of when a bias might occur (during the on-line perception or evaluation phases). Dr. Ste-Marie and I met on several occasions to further solidify the purpose of the study, the methodology, and the potential analyses of the data. After submitting the project for ethical approval, I began writing the thesis proposal and submitted drafts to Dr. Ste-Marie for comments and revisions. The thesis proposal was then submitted and presented to Drs. Pierre Mercier and Glen Kenny. Following their suggestions, minor revisions were made to the literature review, methodology and statistical analysis.

The data was then collected, and Dr. Ste-Marie assisted me in selecting the appropriate data to analyse as well as the appropriate statistical tests. The data was then interpreted and the thesis document was created. Dr. Ste-Marie assisted in the writing process by reviewing and editing all documents. The process of writing the article and elaborated discussion began in April 2001 and was completed in July 2001.

Appendix B

Post-Judgement Questionnaire

Research Questionnaire on the Judging Process

The purpose of this section of the research is to determine whether or not each skater is known within their regional section and/or with you personally. Please respond to each of the following questions for EACH skater to the best of your knowledge based on your personal experiences on the scales provided. Please read all of the questions carefully before you begin to ensure that you respond to each question independently.

1. Do you think that the skater has made a positive name for herself within the Eastern Ontario Section? (I.e., is she known by judges, coaches, officials, etc. to be a good skater?)

	1	2	3	4
	The skater is unknown in the skating community	Some people in the skating community may have heard of this skater	This skater is known by many people in the skating community	This skater is well known
1	Natasha Grela			
2	Rebecca Pivnik			
3	Julie Boone			
4	Sarah Jane Monaghan			
5	Amanda Hunter			
6	Lindsay Myers			
7	Ashton Tessier			
8	Heather Burnett			
9	Christina Nixon			
10	Sabrina Glenn			
11	Jessica Franklin			

	1 The skater is unknown in the skating community	2 Some people in the skating community may have heard of this skater	3 This skater is known by many people in the skating community	4 This skater is well known
1	Natasha Grela			
12	Cheryl McNeil			
13	Sarah Czwnar			
14	Veronica Saikin			
15	Sara Morand			

2. Please rank your personal familiarity with the skaters below.

	NAME	1 No personal contact	2 Minimal personal contact (one or two times)	3 Personal contact on a few occasions	4 Regular interaction (e.g., monitor)
1	Natasha Grela				
2	Rebecca Pivnik				
3	Julie Boone				
4	Sarah Jane Monaghan				
5	Amanda Hunter				
6	Lindsay Myers				
7	Ashton Tessier				
8	Heather Burnett				
9	Christina Nixon				
10	Sabrina Glenny				
11	Jessica Franklin				
12	Cheryl McNeil				
13	Sarah Czwnar				
14	Veronica Saikin				
15	Sara Morand				

3. Before today, how many times do you remember having judged this particular skater over their skating career?

	NAME	1 0 - 2 times	2 3 - 5 times	3 6 - 9 times	4 10+ times
1	Natasha Grela				
2	Rebecca Pivnik				
3	Julie Boone				
4	Sarah Jane Monaghan				
5	Amanda Hunter				
6	Lindsay Myers				
7	Ashton Tessier				
8	Heather Burnett				
9	Christina Nixon				
10	Sabrina Glenn				
11	Jessica Franklin				
12	Cheryl McNeil				
13	Sarah Czwnar				
14	Veronica Saikin				
15	Sara Morand				

Questionnaire de recherche sur le processus de juger

Le but de cette section de la recherche est de déterminer comment chaque patineuse est connue à l'intérieur de son secteur régional et avec vous personnellement. S'il vous plaît répondre aux questions suivantes pour CHAQUE patineuse le mieux que vous pouvez, selon vos connaissances de celle-ci et basé vos réponses selon l'échelle qui vous est fourmit. S'il vous plaît lire attentivement toutes les questions avant de débiter le questionnaire afin d'assurer que vous répondez à chaque question individuellement.

- 1 Est-ce que vous sentez que la patineuse a fait un nom pour elle-même à l'intérieur de son secteur régional? (e.g., jusqu'à quel point est-ce qu'elle est connue d'être bonne par les juges, les entraîneur(e)s, et les officiel(le)s, etc.)

	1	2	3	4
Nom	La patineuse est inconnu(e) dans la communauté du patinage	Certaines personnes de la communauté du patinage ont entendu de cette patineuse	Cette patineuse est connue par la majorité des personnes de la communauté du patinage	Cette patineuse est <u>bien connue</u>
1	Natasha Grela			
2	Veronique Aubin			
3	Cynthia Belleau			
4	Amélie Dagenais			
5	Myriam Chiasson			
6	Sara Bélanger			
7	Josée Charette			
8	Stephanie Thibeault			
9	Karine Fournel			
10	Maryse Larouche			

11	Josee-Lise Courchesne								
12	Jenny-Lynn Armstrong								
13	Maryse Roy								
14	Mélissa Turgeon								
15	Annie-Pier Bertrand								

2. S'il vous plaît placer un crochet dans la case qui décrit le mieux votre familiarité personnelle de chaque patineuse.

	Nom	1 Aucun contact personnel	2 Contact personnel minimale (une ou deux fois)	4 Contact personnel À plusieurs occasions	5 Interaction de façon régulière avec la patineuse (e.g., moniteur)
1	Natasha Grela				
2	Veronique Aubin				
3	Cynthia Belleau				
4	Amélie Dagenais				
5	Myriam Chiasson				
6	Sara Bélanger				
7	Josée Charette				
8	Stephanie Thibeault				
9	Karine Fournel				
10	Maryse Larouche				
11	Josee-Lise Courchesne				
12	Jenny-Lynn Armstrong				
13	Maryse Roy				
14	Mélissa Turgeon				
15	Annie-Pier Bertrand				

3. Avant aujourd'hui, combien de fois rappelez-vous avoir jugé cette patineuse particulière au long de sa carrière de patineuse?

	Nom	¹ 0 - 2 fois	² 3 - 5 fois	³ 6 - 9 fois	⁴ 10 + fois
1	Natasha Grella				
2	Veronique Aubin				
3	Cynthia Belleau				
4	Amélie Dagenais				
5	Myriam Chiasson				
6	Sara Bélanger				
7	Josée Charette				
8	Stephanie Thibeault				
9	Karine Fournel				
10	Maryse Larouche				
11	Josee-Lise Courchesne				
12	Jenny-Lynn Armstrong				
13	Maryse Roy				
14	Mélissa Turgeon				
15	Annie-Pier Bertrand				

Appendix C

Skate Canada Scoresheet (sample)

CANADIAN FIGURE SKATING ASSOCIATION

COMPETITION _____ EVENT _____ DATE _____

REFEREE _____ DIVISION _____

JUDGE'S NO. & NAME		Base Mark	Required Elements	Presentation	TOTAL MARK	O R D I N A L	REQUIRED ELEMENTS						
							Axel Or Double Axel	Double Lutz Jump	Jump Combination	Flying Camel Spin	Combination Spin	Ladies - Layback or Sideways Lean Men - Crossfoot Spin	Circular Step Sequence
Time 2 min 10 sec	Group	Total Deductions					0.1-0.4 (0.5)	0.1-0.4 (0.5)	0.1-0.4 (0.5)	0.1-0.4 (0.5)	0.1-0.4 (0.5)	0.1-0.3 (0.4)	0.1-0.2
1 Natasha Grela													
2 Rebecca Pivnik													
3 Julie Boone													
4 Sarah-Jane Monaghan													
5 Amanda Hunter													
6 Lindsay ^{Maloney} Maloney													
7 Ashton Tessier													
8 Heather Burnett													
9 Christina Nixon													
10 Sabrina Glenn													
11 Jessica Franklin													
12 Cheryl McNeil													

Appendix D
Instruction Set

Instructions #1:

This study involves two stages. First, you will be asked to judge a flight of Novice Ladies' Short Programs from a video tape. The required elements are all from the 2000 skating season (last year), and you will be given a Skate Canada judging sheet which outlines the Group 3 elements. The second stage involves filling out three brief questionnaires about the skaters whom you will be judging. Remember that this is a confidential process and that your marks will not be seen by anyone but myself and research team.

Now I will give you the instructions for the first tape. On the following tape, you will see 15 skaters perform a short program. These Eastern Ontario/*Quebec* skaters were selected from invitational competitions during the 2000 season. After the first skater I will give you the base mark which has been created by expert raters prior to this testing. This step is necessary since we do not have a panel to create a base mark here.

After being given the base mark, you will then be asked to judge each of the 14 skaters that are to follow in terms of the appropriate deductions for errors, the technical merit score, and the artistic impression mark. I would also ask that you keep track of all elements in addition to the marks on the sheet provided. Please do not worry about blocking in or out any of the skaters, just score the programs based on their own merit. Also, do not worry about your handwriting or the short form you use for your notes, just write things in your personal style the way you would in any regular competition. I will go over your short hand with you after you have scored all of the skaters so that I can decipher the marks. All skaters will be presented from approximately the same angle, although due to the videotaping there may be some slight variation. In sum, I would like you to record the elements, deductions, technical, and artistic marks for each skater. Do you have any questions before we begin?

Instructions #2:

Now that you have seen all of the skaters, I will be asking you to fill out three short questionnaires which refer to the skaters you have just judged. The purpose of the questionnaires is, a) to determine whether or not you feel this skater has a positive reputation, based on their past competitive experiences, their potential, any competitive achievements, etc., and b) to determine how many times you recall having judged each skater. If you have had personal contact with a skater, for example, as a monitor, please write that down in the notes section of the form.

To assist you, I will show you a video clip to help you remember the skater, and then I would like you to fill out each question for that particular skater. For example, I will show you Natasha Grella and I would like you to answer the questionnaire on her. I will then show you a clip of the second skater, which you will answer the questionnaire about, then the third and so on. Do you have any questions before I show you the first skater?

Appendix E

List of Deductions for Errors in Novice Ladies Short Program

Technical Deductions

Jumps:

Single Jumps

- 0.1 - 0.4
- 0.4 Fall
- 0.3 Take-off/Landing on two feet
- 0.3 Step out
- 0.2 - 0.3 Rotation cheated
- 0.1 - 0.3 Wrong edge take-off
- 0.1 - 0.3 Insufficient steps
- 0.1 - 0.3 Touch down
- 0.1 Only one preceding move

Combo

- 0.1 - 0.4
- 0.4 Fall on take off or first jump
- 0.4 Both jumps landed on two feet
- 0.3 Step out of first jump
- 0.3 Fall on second jump
- 0.3 less than required rotation (either jump)
- 0.2 Step out on second jump
- 0.2 Two foot on either jump
- 0.1 - 0.3 Wrong edge take-off
- 0.2 Either jump rotation cheated
- 0.2 One jump landed on two feet
- 0.1 - 0.3 Touch down
- 0.2 Three turns between the jumps

Spins

Flying Spin

- 0.1 - 0.4
- 0.4 Fall
- 0.1 - 0.3 Position in the air not attained
- 0.1 - 0.3 Less than required rotations
- 0.1 - 0.2 Incorrect takeoff or landing
- 0.1 - 0.2 Touch down

Solo Spin 0.1 - 0.4

- 0.4 Fall
- 0.1 - 0.3 Less than required rotations
- 0.1 - 0.2 Incorrect takeoff or landing
- 0.1 - 0.2 Touch down

Combo 0.1 - 0.4

- 0.4 Fall
- 0.1 - 0.3 Position in the air not attained
- 0.1 - 0.3 Less than required rotations
- 0.1 - 0.2 Incorrect takeoff or landing
- 0.1 - 0.2 Touch down

Step/Spiral Sequence

- 0.1 - 0.3
- 0.3 Fall
- 0.2 Pattern incorrect
- 0.2 Jump of more than 1/2 rotation
- 0.2 No change of foot (spiral only)
- 0.1 - 0.2 Stumble
- 0.1 - 0.2 Retrogression
- 0.1 - 0.2 Less than three spirals

Other

- 0.3 repetition of jump
- 0.1 - 0.2 Not according to regulations
- 0.1 - 0.2 Extra or repeated element
- 0.1 Illegal move (deduct from both marks)

Appendix F

Consent Form and Information Sheet

CONSENT AND INFORMATION FORM

This research is being conducted by the following persons. If for any reason you would like to contact one of us, the following information should provide you with the means to do so.

Dr. Diane Ste-Marie
 Associate Professor
 School of Human Kinetics
 University of Ottawa
 Phone Number: 562-5800 ext. 4255

Leanne Findlay
 Graduate Student
 School of Human Kinetics
 University of Ottawa
 Phone Number: 562-5800 ext. 4248

I, _____, am interested in collaborating in the research conducted by Leanne Findlay, under the supervision of Professor Diane Ste-Marie of the Department of Health Sciences, School of Human Kinetics at the University of Ottawa. The purpose of the research is to identify the cognitive processes underlying the performance appraisal process in figure skating judging.

My participation will consist of attending one session lasting approximately one hour. During this period, I will be asked to judge fifteen Novice Ladies' short programs and to fill out a brief questionnaire on these skaters. The sessions have been scheduled for _____. I will be asked to judge the skaters based on technical merit, artistic marks, and the appropriate deductions for each program. I understand that the contents will be used only for the purpose of understanding the judging process and the factors which influence this process. I am aware that my confidentiality will be respected since all data will be viewed by the researcher and her advisor exclusively, and all data will be kept in a locked and restricted area, to which only the researcher, her supervisor and assistants have access. After three years post-publication, all raw data will be destroyed. My anonymity will be maintained by assigning my scoresheet a number rather than my name, and all data will be published only in pooled format. This means that my data will be compiled with all other participants and no individual responses will be identified in any way. Also, my name and identity will be completely concealed, and at no point will I be identified as a participant.

I understand that since this activity deals with information from my judging marks, it may cause me to feel some psychological discomfort when making a subjective evaluation of skaters. I have received assurance from the researchers that every effort will be made to minimize any discomfort. I am also aware that there may be some deception involved in the task, meaning that I will not be fully aware of the reason for doing certain tasks. If, however, deception does occur, I will be fully debriefed following the judging task. If at any time I feel uncomfortable, I am aware that I am able to discontinue participation. Upon the completion of the task, the researcher will fully debrief me as to the data analysis and I always have the freedom to ask that my data be excluded without prejudice. Prior to publication, I have the option to receive a copy of the results so that I am aware of the findings in advance of publication.

Any information requests or complaints about the ethical conduct of the project may be addressed to the relevant University Research Ethics Board for Health Sciences, or by calling the Protocol Officer for Ethics in Research, Lise Frigault, Room 302, Tabaret Hall, 562-5800 ext. 1787. Also, there are two copies of the consent form, one of which I may keep in case I have any further questions or concerns.

Researcher's signature: _____ Date: _____

Research Subject's signature: _____ Date: _____

FORMULAIRE DE CONSENTEMENT

Dr Diane Ste-Marie
 Professeure Associée
 École des Sciences de l'Activité Physique
 L'Université d'Ottawa
 Numéro de téléphone: 562-5800 ext. 4255
dstmarie@uottawa.ca

Leanne Findlay
 Étudiante Graduer
 École des Sciences de l'Activité Physique
 L'Université d'Ottawa
 Numéro de téléphone: 562-5800 ext. 4248

Je, _____, suis intéresser à collaborer à la recherche diriger par Leanne Findlay sous la supervision de professeur Ste-Marie de l'École D'Activité Physique, Faculté des Sciences de la Santé de l'Université D'Ottawa, où le but de cette recherche est d'identifier les processus cognitifs lors des jugements de compétitions de patinage artistique.

Ma participation consistera à une séance d'une heure durant laquelle je jugerai quinze femmes Novices, et je répondrai ensuite à un questionnaire sur ces patineuses. La session a été planifiée pour _____. J'aurai à jugé les patineuses sur: le mérite technical, le mérite artistique et les déductions mérité pour chaque programmes. Je comprends que le matériel recueilli sera utilisé pour comprendre le processus de jugement et les facteurs qui influencent ce processus. Toute donnée amassée au cours de cette session sera utilisée uniquement à fin de recherche et de publication. J'ai reçu l'assurance du recherchiste que les renseignements fournis seront strictement confidentiels. L'anonymat sera assuré de la façon suivante: mon nom ne sera pas imprimé dans aucun matériel de collection de données et, ainsi que tous les filières de données seront assigner des numéros et non pas un nom. Toute donnée qui sera publiée sera sous forme de groupe. De plus, mon nom et mon identité restera concilier et en aucun temps je ne serai pas identifier comme participant. Les données seront détruites trois ans après la publication, tel que requis par les publications de recherche

Je comprends que cette activité est reliée à l'information sur mes évaluations, cela pourrait me faire sentir inconfortable psychologiquement lors de mes évaluations subjectives des patineuses. J'ai reçu l'assurance des chercheurs que tout les efforts ont été mit en place pour minimiser ces malaises. Je comprends que je pourrais ressentir un certain niveau de déception lors de la tâche, la raison étant que je suis conscient des raisons que je performe certaines tâches. Si la déception a lieu, je serai informé après la session d'évaluation. Je suis libre de me retirer de la recherche en tout temps, avant ou pendant l'expérience, refuser de participer à certaines phases de recherche sans encourir de préjudice sous aucune forme.

Tout renseignement requis ou plaintes concernant la conduite éthique de ce projet peuvent être adressé au Comité d'éthique de la recherche en sciences de la santé et sciences ou en appelant la responsable de la déontologie en recherche: Lise Frigault, appartement 302, Pavillon Tabaret, 562-5800 poste 1787.

Signature du chercheur: _____ Date : _____
 Signature du participant: _____ Date : _____

Appendix G
Debriefing Sheet

DEBRIEFING FOR REPUTATION BIAS IN FIGURE SKATING

We would now like to thank-you for having participated in our research and explain exactly what we are investigating. First, we will provide some background information for our research questions. We will then describe in more detail the experiment which you participated in, and finally we will review our methods for analysing the data. If, upon reading this form, you feel at all uncomfortable about your participation, please feel free to question the researcher or to request to exclude your own data from any analysis. Also, please be assured that this testing is by no means intended to evaluate your ability or competency as a judge, it is merely being done to observe trends across a group of judges. For this reason, all data will be published in group format, with no names or individual results being open for publication.

Research in the field of subjective evaluation has focussed on the effects of expectations on judging. Scheer & Ansoerge (1980) have found in gymnastics judging that judges are often influenced by what is termed a "within-team order effect", whereby judges tend to give higher marks to athletes who perform later in a position of rotation than those who are in an earlier position. They argued that the reason for this occurrence is that judges have the **expectation** that coaches choose their best athletes to compete last, thus marks are reserved for the expected-to-be-better performances. Our interest is whether or not an expectation for performance occurs as a result of a skater's reputation within the skating community. More specifically, we question whether athletes who are well known will receive higher marks based on their name rather than performance related factors. Considering that the range of technical deductions for various elements as well as the subjectiveness of evaluation, it is important to identify any expectation effects that may exist. Our hypotheses were that a judges would allocate higher technical and artistic marks and lower deductions to athletes whom they felt were known to have a good reputation.

In order to investigate this question, two formats of the same 15 programs on videotape were generated. You have seen a tape which identified seven skaters from your home section with their name. The other seven skaters were from the Quebec (QC) section and were identified by a fictitious name. The opposite was done for the Quebec judges. In this way, we are creating two groups: known athletes (home section) and unknown athletes (out of section) for judges from the two different sections.

Our second research question is with respect to cognitive processing in judging. As identified on the consent form, we are interested in the performance appraisal process, and we will investigate this process based on the deductions versus the technical and artistic marks. We are hoping to assess current models of the evaluation process and apply our findings to a social cognitive framework. To our knowledge, this has not yet been done for sport evaluation.

In terms of the application of this research, we hope to answer the popular question as to whether or not expectations based on an athlete's name in sport influences the performance appraisal process. We are hoping to use this information to make recommendations regarding the evaluation process in figure skating. Again I remind you that in no way are we evaluating your performance as a judge, we are merely assessing the general process of figure skating judging and the way (unconsciously or consciously) that judges are affected by the reputation of an athlete. If you have any questions or comments, feel free to ask now or to contact the researchers listed on the consent form. Again, thank you for participating in this research.

Leanne Findlay
Graduate Student

Diane Ste-Marie
Supervising Professor

Compte rendu de l'influence du biais de la réputation en patin artistique

Nous aimerions à ce moment vous remercier de votre participation à notre recherche et aussi vous expliquer exactement le but de celle-ci. Nous allons premièrement vous procurer un bref historique de notre problème de recherche. Ensuite, nous décrirons, de façon plus détaillée, l'expérience à laquelle vous venez de participer. Finalement, nous allons résumer les méthodes d'analyse des données utilisées. Si à n'importe quel moment durant la lecture de ce document vous vous sentez inconfortables d'avoir participé à cette expérience, s'il vous plaît en parler au chercheur ou demandez d'avoir vos résultats exclus des analyses subséquentes. Aussi, veuillez noter que ce test n'est sous aucune circonstance une évaluation de vos habiletés ou compétences à juger - c'est uniquement une observation des tendances dans les différents groupes de juges. Pour cette même raison, les résultats seront publiés suivant un format de groupe, ne contenant aucun nom ou résultat individuel dans la publication.

La recherche dans le domaine de l'évaluation subjective a surtout étudié les effets des attentes dans le jugement d'événements sportifs. Scheer & Ansorge (1980) ont découvert que les juges sont souvent influencés par ce qu'on appelle "within-team order effect", une tendance à donner des scores plus élevés aux athlètes qui se présentent plus tard dans une rotation donnée comparativement à ceux qui passent plus tôt. Ils argumentent que ce phénomène résulte des attentes que les juges ont de l'ordre de présentation des athlètes. Les juges croient que les entraîneurs décident de présenter leurs meilleurs athlètes à la fin et les scores sont donc réservés pour les performances à venir. Notre intérêt particulier dans le domaine est de découvrir si l'attente de performance existe en raison de la réputation d'un athlète donnée dans le cadre du patinage artistique. Plus précisément, nous voulons savoir si les patineurs qui sont bien connus recevront des scores plus élevés basés sur leur réputation plutôt que sur des facteurs reliés à la performance. Prenant en considération les multiples déductions techniques pour différents éléments en plus de la subjectivité de l'évaluation en général, il est important d'identifier quelque attente qui puisse exister. Notre hypothèse était que les juges alloueraient des scores techniques et artistiques plus hauts et des déductions techniques plus basses aux athlètes connus et qui, d'après ces juges, avaient une bonne réputation.

Afin de pouvoir faire notre enquête, deux versions de bande vidéo ont été formatées, contenant chacune les mêmes 15 programmes. Vous avez visionné la bande vidéo qui identifiait, par nom et club de patinage artistique, sept patineurs de votre région et aussi sept patineurs venant de l'Est de l'Ontario, présentés avec des noms fictifs. La bande vidéo opposée fût présentée aux juges de l'Ontario. De cette manière, nous avons réussi à simuler deux groupes; les athlètes connus (région propre) et les athlètes inconnus (autre région) pour les juges des deux différentes régions.

Notre deuxième problème de recherche est relatif aux processus cognitifs du jugement sportif. Comme nous l'avons déjà mentionné sur le formulaire de consentement, nous sommes intéressés par le processus de l'évaluation de la performance. Nous étudierons ce phénomène en comparant les déductions en comparaison avec les scores artistiques et techniques. Nous espérons ainsi faire l'évaluation des modèles actuels du processus d'évaluation et de mettre nos résultats en pratique dans un cadre de travail socio-cognitif. Ceci est un projet qui, à notre connaissance, n'a pas encore été fait en ce qui concerne l'évaluation sportive.

En ce qui a effet aux modalités de cette recherche, nous espérons répondre à la question populaire de l'influence de la réputation d'un athlète sur les attentes et ainsi l'évaluation des juges. Notre intention est de pouvoir utiliser cette information pour formuler des recommandations à l'égard du processus d'évaluation en patinage artistique. J'aimerais vous rappeler encore une fois que ceci n'est pas une évaluation de votre capacité en tant que juge, mais simplement une évaluation du processus général d'évaluation en patinage artistique et de la manière dont les juges sont influencés, de façon conscientes ou inconscientes, par la réputation des athlètes. Si vous avez des questions ou des commentaires, prenez s'il vous plaît le temps de nous adresser ou de contacter les chercheurs dont le nom apparaît sur le formulaire de consentement. Nous vous remercions encore une fois de votre participation.

Leanne Findlay
Étudiante de deuxième cycle

Diane Ste-Marie
Professeur supervisant le projet de maîtrise

Appendix H
Ethics Approval



Université d'Ottawa • University of Ottawa

Cabinet du vice-recteur
à la recherche

Office of the Vice-Rector,
Research

HEALTH SCIENCES AND SCIENCE RESEARCH ETHICS BOARD

CERTIFICATION OF ETHICAL APPROVAL

This is to certify that the University of Ottawa Health Sciences and Science Research Ethics Board has examined the Application for Ethical Approval for the research project The Name Game: Is there a Reputation Bias in Figure Skating Judging? (File H05-00-05) submitted by Leanne Findlay. The REB found that this project meets appropriate ethical standards as outlined in the Tri-Council Policy Statement and in the Procedures of the University of Ottawa Research Ethics Boards and accordingly gave it a Category Ia (Approval). This certification is valid for one year from the date indicated below.

Lise Frigault

Protocol Officer for Ethics in Research,
for the Chairperson of the Health Sciences and Science REB
Valerie Whiffen

AUGUST 31, 2000

Date

FILE NUMBER: H05-00-05