## Articles

# Measuring the Multidimensional Nature of Sporting Event Performance Consumption 

Robert Madrigal<br>University of Oregon


#### Abstract

Sporting events are a type of unscripted skill performance that are bound by a set of rules and are contested so as to produce an outcome wherein one performer can claim an unequivocal victory over another. Although a predominant form of leisure behavior worldwide, little is known about what sports spectators attend to while consuming such performances. Conceptualized here as a multidimensional concept, this article details the development of a reliable and valid measurement scale of sporting event consumption that is represented by two higher-order factors (Autotelism, Appreciation) that are each comprised of three unidimensional factors (Fantasy, Flow, Evaluation; Personalities, Physical Attractiveness, Aesthetics, respectively). Collectively, the scale is referred to as FANDIM. Differences across FANDIM dimensions according to sports group (i.e., aesthetics vs. purposive) and respondent sex are investigated.


KEYWORDS: Sports fans; performance consumption.

## Introduction

The consumption of leisure activities is experiential in nature and involves absorbing the symbolic meanings associated with more subjective characteristics. The consumption phenomena itself is multifaceted and its pleasures intrinsic. The current paper investigates the consumption of a skill performance which is defined as one that is witnessed by an audience either directly or indirectly via media, and whose outcome relies on the abilities of the actors charged with delivering the performance (Deighton, 1992). Contests are a type of skill performance in which competitors seek to demonstrate excellence in the hope of attaining a favorable outcome (Barthes, 1972). The type of contested skill performance considered here is the consumption of competitive sporting events. In 2001, Americans spent an estimated $\$ 26$ billion on attending sporting events and North American companies invested nearly $\$ 34$ billion in sports-based sponsorships and advertising (Sports Business Journal, December 20, 2002). Yet, in spite of its

[^0]economic impact, little is known about the dimensions underlying the experience of sporting event consumption. Most work on the topic has either been conceptual in nature (Sloan, 1989; Zillmann \& Paulus, 1993) or focused on spectators' motives for watching sports (Gantz and Wenner, 1991; Wann, 1995) rather than on the consumption experience itself. Madrigal (2003) provided a description of the antecedents and consequences of a live sporting event as it transpired, but did not address those specific aspects of sporting events that are attended to by fans during consumption.

The purpose of this study is twofold. First, a general definition of skill performance consumption is provided within a broader nomological network. A brief discussion of related yet different theoretical constructs is also offered. The discussion is then narrowed so as to focus on the specific context of skill performance considered here, sporting events. Second, the article outlines the development of a set of parsimonious scales, henceforth referred to collectively as FANDIM, designed to measure the underlying dimensions of sporting event consumption.

## The Consumption of Skill Performance

Deighton (1992) described skill performances such as sporting events or jury trials as staged displays of competence occurring in naturalistic settings that emphasize the event's realism. This differs from a show performance (e.g., theatre) that is contrived for the audience's benefit, occurs in an artificial setting and emphasizes elements of fantasy. Although the role of the observer in both types of performance is as a witness to the action rather than as a direct participant, the outcome of a show performance is usually predictable or ritualistic whereas skill performance is characterized by tension and uncertainty about the eventual outcome.

Skill performance consumption, therefore, refers to the manner in which a spectator (an attendee or media consumer) interacts with the witnessed action that occurs during an event for which the outcome is uncertain. It is conceptualized as a multifaceted phenomenon that underlies the nature of a spectator's experience. Rather than focusing on the transitional and evolving affective states likely to arise during a performance (see Madrigal, 1995, 2003), the perspective used here considers the dimensions underlying consumption to be enduring. A spectator who appreciates the aesthetic quality of a gymnastics meet, for instance, should attend to this facet of consumption regardless of her emotional arousal while watching. It is possible, however, that a spectator's appraisal of aesthetic quality may contribute to specific emotions.

## Related Concepts

The facets of a skill performance attended to by a spectator during consumption are thought to be different than affective reactions. So too are they thought to be different than involvement which refers to a person's
attention to some object because of its perceived relevance or importance (see Havitz \& Dimanche, 1999 for a review). What sets the multidimensional conceptualization of involvement (Laurent \& Kapferer, 1985) apart from what is being proposed here is that involvement focuses on the motives underlying a person's interest in a product class. Motives induce people to behave in a certain way. Thus, the multiple dimensions of involvement reflect certain needs that are being met by a product or product class which, in turn, influence subsequent purchase behavior. In contrast, the dimensions described in this paper are concerned with those aspects of a skill performance that are attended to by spectators during consumption, not spectator needs being met by consumption.

Competitive sporting events represent a type of contested skill performance (Barthes, 1972). Although external factors related to the presence of others are likely to make the experiences qualitatively different, the underlying factors describing consumption in both cases are thought to be the same. The underlying dimensions attended to by sporting event spectators are thought to be consistent across various contests in a particular sport, but are likely to vary in importance across different types of sports. Moreover, it is expected that differences across the various dimensions of consumer experience will exist between spectators of the same sport.

Sporting event consumption is related to, but distinct from, the motives for watching sports and team identification. For the most part, motives have been conceptualized as predictors of a person's interest in watching sports generally, rather than interest in watching a particular sport. Sloan (1989) discussed a number of theories to explain people's motives for watching sports events (see also Wann, Melnick, Russell, \& Pease, 2001 for a review). Interestingly, most motives have focused on a spectator's general interest in watching sports rather than focusing on a specific type of sport (e.g., team vs. individual) or specific sports.

Motives for watching sports have been shown to vary by the viewer's sex. Gantz and Wenner (1991) found that compared to men, women were more likely to watch sports for social reasons, whereas men watched in order to get psyched up, relax, let off steam, and drink alcohol. Men were also motivated to watch because they enjoyed the tension and drama of a competitive contest and because it gave them something to talk about. Research by Wann (1995) and his associates (Wann, Schrader, \& Wilson, 1999) revealed that males reported higher levels of the following motives: eustress (i.e., pleasant stress), self-esteem, escape and aesthetics (James \& Ridinger, 2002 also found a similar effect for the aesthetics motive). Women, on the other hand, reported higher levels of family motivation in both studies (see also DietzUhler, Harrick, End, \& Jacquemotte, 2000). In sum, motives have been broadly defined and used to measure a spectator's general interest in watching sporting events.

The topic of team identification has also been well represented in the sports spectator literature. Team identification represents the extent to which a person feels a psychological attachment to a particular sports team (see

Madrigal, 2004 and Wann et al., 2001 for reviews). By basking in the reflected glory of a successful team, spectators enhance their own self esteem and satisfy achievement needs (Cialdini, Borden, Thorne, Walker, Freeman, \& Sloan, 1976). Zillmann, Bryant, and Sapolsky (1989) viewed alliance formation with a team in terms of disposition theory and noted that spectators' favorable/unfavorable dispositions toward competing teams serve as the basis for their emotional reactions. Accordingly, meeting the need for enjoyment requires spectators to be favorably disposed toward one of the competing teams. It would appear from the research in this area, therefore, that team identification and disposition toward the team represent means by which specific motives (e.g., self-esteem, achievement seeking) for watching sports are satisfied.

## The Dimensions of Sporting Event Consumption

In contrast to a motive-based model intended to satisfy specific needs, the current study considers the dimensions along which sporting events are consumed. Conceptualized as a multifaceted concept, sporting event consumption is hypothesized to consist of two higher-order dimensions-each comprised of multiple latent first-order factors. The factors represent the specific facets of consumer experience with each operationalized as a unidimensional set of individual items. The two higher-order factors were derived from the literature on experiential consumption and skill performance. Specifically, skill performance consumption is thought here to reflect experiential elements related to (a) autotelism (Csikszentmihalyi, 1975; Holt, 1995), aspects of consumption that have a purpose in and not apart from themselves; and (b) appreciation (Holt, 1995), elements that tap the performance event's situations, people, and actions. The unidimensional constructs of flow, fantasy and evaluation are considered to be autotelic because each is an element of the consumption experience that is an end unto itself and reflects the consumer's immersion in the event. In contrast, the factors of aesthetics, personalities and physical attraction each involve an appreciation or estimation of the qualities inherent in the athlete and the sport. Each dimension will now be discussed.

Autotelism. Csikszentmihalyi (1975) described autotelic experience as "a psychological state, based on concrete feedback, which acts as a reward in that it produces continuing behavior in the absence of other rewards" (p. 23). Autotelic experience focuses on the consumer's vicarious interaction with the performance and is thought to underlie three FANDIM dimensions. Flow represents an optimal psychological state characterized by intense absorption, a loss of self consciousness, and an altered sense of time (Csikszentmihalyi, 1990). Fantasy, on the other hand, evokes a sense of playfulness typified by feelings of escape, pleasure, and relaxation. In contrast to commercially created fantasy settings such as Las Vegas (see Shields, 1992) or participatory fantasies (Belk \& Costas, 1998), the focus here is on how sporting event consumers might fantasize that they are part of the action (see Hirschman and Holbrook, 1982).

The final autotelic element proposed here is evaluation. In an ethnography examining how consumers consume a baseball game, Holt (1995) noted that spectators use a variety of norms and baseline expectations developed from their observation of the game over time to evaluate baseball. They make value judgments about the quality of play and assess the actions of the performers responsible for delivering the performance (see also Deighton, 1992). Thus, evaluating is defined here as the act of critically judging both the quality of the action and the athletes' skills in their quest to achieve a desirable outcome (i.e., win). It is an aspect of autotelic experience because passing judgment on the action reflects an ongoing interaction with the event that elicits greater evaluation. There is no extrinsic reward for engaging in this aspect of consumption. Instead, it is an act that leads to a deeper involvement with the event itself.

Appreciation. Rather than focusing on the interactive elements of a sports performance, appreciation emphasizes spectators' consumption of the artistry displayed in the sport itself and the personal characteristics of those involved in delivering the performance. Three dimensions are considered. The first, aesthetics, refers to consumers' appreciation for the grace and beauty of the sport itself. Aesthetics represent a response to the mastery exhibited by the athletes executing the action (Boxill, 1985; Guttmann, 1986; Holbrook \& Zirlin, 1985; Holt, 1995) and is often used as a basis for determining event outcomes in certain sports such as gymnastics and springboard diving (Best, 1978; Zillmann and Paulus, 1993).

Also appreciated in a skill performance are the unique personalities of the athletes. Appreciation of the performers' personalities refers to the act of focusing attention on specific athletes during a contest, usually those possessing the greatest prominence, renown, or notoriety. For many, appreciating a specific athlete's personality often takes precedence over appreciating the sport itself. Witness the $37 \%$ drop in television ratings (i.e., 16 million viewers) for the 1999 National Basketball Association Finals compared to those of the previous year in which Michael Jordan played (Sports Business Daily, June 29, 1999).

Sporting event consumers are also thought to appreciate the physical attractiveness of the competitors. Just as ancient cultures widely recognized and even celebrated the sensual appeal of the athletes competing in sporting events (Guttmann, 1996), modern fans also frequently engage in a form of voyeurism while consuming spectator sports (Duncan \& Brummett, 1989; Hofacre, 1994; Morse, 1983; Mulvey, 1975). References to this dimension have also appeared in a variety of popular media such as Sports Illustrated (Silver, 1997), Nightline (August 26, 2002: "Game, Sex, Match"), and the San Francisco Chronicle (Ganahl, 2004).

## Methodology

The remainder of the article reports the development and validation of a measurement instrument designed to assess the proposed dimensions along which sporting event performances are consumed. It begins with a
description of the procedure used to generate items and establish content validity. Next, data from the first study are used to describe how specific items were selected. Four subsequent studies investigating factorial invariance, scale reliability and multiple forms of validity are discussed. Next, tests investigating subscale differences based on sex and sport type are presented, as well as tests considering the relations between each dimension and interest in watching the sport for which the FANDIM scale was completed. Data from Study 4 are then used to compare the dimensions of the FANDIM instrument to other known measures in order to establish concurrent validity. Study 5 then reports the results of a test-retest procedure. Finally, a general discussion is presented along with study limitations and a conclusion.

## Item Generation and Content Validity

Content validity is concerned with the extent to which a specific set of items reflects a "randomly chosen subset of the universe of appropriate items" (DeVellis, 1991, pp. 43-44). Initially, three focus groups were conducted in which small groups ( $n<12$ ) of students and non-students were asked to discuss what they attended to and/or thought about while watching various types of competitive sporting events. Frequently mentioned points were converted to statements and used in the initial item pool. Multiple thesauri were then consulted, as well as existing scales on sport spectator motives (e.g., Gantz \& Wenner, 1991; Hansen \& Gauthier, 1989; Kahle, Kambara, \& Rose, 1996; Wann, 1995) and the flow experience (Jackson \& Marsh, 1996). An initial set of statements was then constructed to reflect those elements of the experience attended to by viewers while watching sporting events. Item screening designed to eliminate redundant, double-barreled, ambiguous, and leading statements led to an initial pool of 68 items. Each item was then grouped into one of the six dimensions by the researcher.

Content validity is enhanced when a panel of experts reviews each item for its relevance to a proposed domain (DeVellis, 1991). Eleven individuals who had previously published research on sports spectator behavior or on the underlying psychological processes describing fan behavior were asked to act as judges. Nine agreed to participate. Each expert was given a concise definition of three of the six dimensions and its respective items. Each dimension was evaluated by at least four and no more than five judges. The judge was asked to identify the six items that best represented the given dimension's definition. Judges were also given the option of recommending additional items or changing an item's wording. Items ranked in the top six by at least half of the judges examining a dimension were retained. No items were volunteered by a judge. The review yielded 36 items with each dimension represented by five to eight items.

## Study 1: Item Purification

Method. Data for item refinement were collected in classroom settings from 610 college-aged students attending one of six North American universities. Four versions of the scale were developed, each based on a random
ordering of the items. Using a 10 -point scale ranging from not at all like me (1) to very much like me (10), respondents were asked to indicate how descriptive each item was of their thoughts while watching their favorite sport.

Item analyses. Respondents were randomly split into two groups. The items comprising each of the six dimensions were then submitted to a separate principal-axis factor analysis for each group. In total, twelve separate factor analyses were performed (i.e., one analysis for each of the six dimensions for each sample). For each sample, a Cronbach's alpha was also estimated for the set of items comprising each dimension. The initial criteria used to retain an item were that it must (a) have an average (i.e., across both samples) factor loading of .60 or better; (b) not cross load at greater than .20 on any other factor emerging from the analysis in either sample; and (c) have a corrected average item-to-total correlation above .40. Items meeting these criteria were then entered into a single factor analysis performed on each factor using the entire sample of 610. In order to ensure parsimony, only the top four items loading on a factor were retained.

## Studies 2-4: Confirmatory Analyses and Tests of Factorial Invariance

Method. As in Study 1, respondents in each subsequent study were asked to indicate how descriptive each item was of their thoughts while watching the selected sport ( 10 -point scale, not at all like me/very much like me). Participants were also asked to indicate: (a) their interest in the sport on a four-point scale (uninvolved, casual interest, fan, die-hard fan) and (b) how likely it was that they would either attend a competition or watch the sport on television (i.e., likelihood of watching; 10-point scale, not at all likely/extremely likely).

Data for Study 2 were collected from 678 students in a classroom setting attending one of two universities and representing a variety of majors. Respondents were presented with one of eight different versions of the scale corresponding to a 4 (Sports Group) $\times 2$ (Item Order) factorial in which the 24 items were initially ordered randomly. ${ }^{1}$ The four sports groups, each consisting of three different sports, were created on the basis of Best's (1978) contention that a sport may be classified as having either a purposive or aesthetic function. Purposive sports are those in which the purpose of the action is independent of the manner in which it is achieved. In baseball, for example, the purpose of playing is to score more runs than the opponent. This can be achieved in a variety of ways. Although the action displayed in purposive sports may be judged as having aesthetic appeal, it is incidental to the primary purpose. In contrast, the purpose of aesthetic sports can't be separated from the means by which it is achieved. In fact, these sports are

[^1]explicitly judged on aesthetic norms. The aim of a springboard diver is to not simply fall from the board to the water. Rather, divers are rewarded for the way in which appropriate movements are executed prior to and including entry into the water. Demonstration of the aesthetic is not incidental; it is central to how the performance is scored.

The sports groups were further differentiated based on performer sex. Thus, the four sports groups were: aesthetic sports (figure skating, gymnastics and springboard diving) performed by (1) women and (2) men; and purposive sports performed by (3) women (basketball, softball, volleyball) and (4) men (basketball, football, baseball). Each participant was instructed to indicate which of the three within the assigned sports group was most preferred and to answer all FANDIM items in regard to this spectator sport. Because the study was concerned with how spectators consume a sporting event, it was felt that each respondent should have at least some interest in watching the selected sport. Thus, only those indicating at least a casual interest in the sport were retained for analysis. This yielded a final sample of 624 , in study 2 , most of whom were Caucasian ( $73 \%$ ) and a fairly equal representation of females ( $52 \%$ ) and males.

A professional research firm collected data for Study 3 at shopping malls located in six U.S. cities. Two hundred sixty three respondents were asked to select their most preferred sport from the same groups of three described in Study 2. Again, only those indicating at least a casual interest in the selected sport were asked to participate. Approximately equal numbers of respondents were assigned to each of the sports set conditions. The mean age of the respondents was 35.4 years ( $\mathrm{SD}=11.76$; range $=18-69$ ), slightly over half were women ( $50.7 \%$ ), and over $30 \%$ were college graduates. The ethnic composition of the sample was predominately Caucasian ( $61.9 \%$ ), followed by Hispanic ( $18.5 \%$ ) and African American ( $10.7 \%$ ).

The purpose of Study 4 was to examine the construct validity of the FANDIM scale. Data were collected from 372 undergraduates using the same 4 (Sports Groups) $\times 2$ (Item Order) factorial design that was described above. The majority of respondents were female ( $58 \%$ ), Caucasian ( $73 \%$ ), and in their first year of college ( $46 \%$ ). In addition to completing the FANDIM items, respondents were asked to complete a number of other scales thought to represent concepts that should be theoretically related to one or more of the FANDIM dimensions. The description of each scale is provided in the Results section of the paper.

## Study 5: Test-Retest

Method. For purposes of establishing test-retest reliability, 44 MBA students completed the FANDIM scale in a classroom setting. There was a fiveweek lag between test administrations. Respondents completed the scale in regard to how they consume a football game. All indicated at least a casual interest in the sport. Most respondents in this sample were women ( $61 \%$ ) and Caucasian ( $70 \%$ ) with a mean age of slightly over 26 years.

## Analysis of Competing Models

Given the sensitivity of chi-square tests to sample size, four additional fit indices were used in this research. The root-mean-square error of approximation (RMSEA) is an absolute-fit measure. The confidence interval (CI) surrounding RMSEA suggests that over all randomly sampled point estimates, the true RMSEA will be captured inside the associated interval range $90 \%$ of the time. Thus, a model should not be rejected in favor of an alternative unless the latter's RMSEA point estimate falls outside the $90 \% \mathrm{CI}$ of the former. Also used to assess model fit were the comparative fit index (CFI) and the standardized root mean squared residual (SRMR). Hu and Bentler (1999) recommend the following cutoffs for determining acceptable fit: RMSEA < .06; CFI > .95; and SRMR < .08. Finally, Bozdogan's (1987) consistent version of Akaike's information criterion (CAIC) is included as a parsimony-based index that accounts for statistical goodness of fit relative to the number of parameters needed to be estimated to achieve that fit while also accounting for sample size. The CAIC statistic is used to compare models. Lower CAIC values are preferred because they indicate greater parsimony.

## Results

## Comparison of Competing Models

To assess the adequacy of the proposed higher-order factor structure, confirmatory factor analysis was used to compare a number of alternative models. The models were estimated using LISREL 8.53. ${ }^{2}$ The alternative factor structures tested were as follows: a null model (M1); a one-dimensional model for which all the observed variables were forced to load on a single first-order factor (M2); a two-factor uncorrelated model in which all the observed variables were loaded on either a single Autotelism or single Appreciation first-order factor (M3); a similar two first-order factor correlated model (M4); a six-factor orthogonal model (M5); a six-factor first-order correlated model (M6); a second-order factor model with two higher order factors, Autotelism and Appreciation comprised of three first-order factors each (Fantasy [FAN], Flow [FLO], Evaluation [EVA]; and Personalities [PER], Physical Attractiveness [PHY], Aesthetics [AES], respectively; M7); and an identical model to the latter but with AES cross-loading on both second-order factors (M8; see rationale below).

In general, the two-factor higher-order model (M7) and the correlated six first-order factor model (M6) provided the best pattern of fit statistics for

[^2]the data from each of the three studies. ${ }^{3}$ Based on CAIC value, M7 was preferred over M6 in studies 3 and 4. A review of the modification indices (and subsequent CAIC values) from the Study 2 analysis revealed that allowing AES to cross-load on both second-order factors (M8) resulted in a more parsimonious fit than M6 or M7. ${ }^{4}$ M8 fit statistics were: $\chi^{2}=537.53$, $\mathrm{df}=$ 244, RMSEA $(90 \% \mathrm{CI})=.044(.039-.049), \mathrm{CFI}=.98, \mathrm{SRMR}=.048$.

Convergent validity. The individual items included in the FANDIM scale and a summary of tests related to the internal consistency of the first-order factors are shown in Table 1. The table features the completely standardized loadings of the items on their respective latent constructs; the average variance explained by each construct; coefficient alphas and composite reliabilities for the constructs; and the average variance explained (AVE) by each construct. For each of the samples, the results of these tests suggest that the six subscales were reliable. According to Fornell and Larcker (1981), convergent validity is evidenced when the composite reliability of a factor is equal to or greater than .60 and AVE is equal to or greater than .50 . As indicated in Table 2, composite reliabilities ranged from .90 to .94 across studies with a mean of .92 . The mean AVE was .75 with a range of .68 to .82 . In addition, as shown in Table 2, the first-order factor loadings were all highly significant ( $t$ 's ranged from 13.14 to 34.41 with a mean of 25.84 ). Thus, the scale reliabilities for the first-order factors were deemed satisfactory.

An analysis of the higher-order factor analysis also suggests satisfactory reliability. All first-order factor loadings on their respective second-order factors were substantial and significant ( $t$ values ranged from 3.78 to 18.78 across studies, with a mean loading of 15.74 on the Autotelism dimension and a mean loading of 9.57 on the Appreciation dimension, all $p$ 's $<.001$ ). Correlations between the second-order factors were significant in Studies 2, 3 and $4(\beta$ 's $=.37, .85, .88 ; ~ t ' s=6.22,18.00,15.97$; all $p$ 's $<.001)$.

Discriminant validity. The discriminant validity of the first-order factors was established in three ways. First, for all studies, none of the construct correlations were within two standard errors of unity. Second, the average variance extracted (AVE) for each construct was greater than its squared correlation with any other construct (Fornell and Larcker, 1981). Finally, a series of chi square difference tests comparing the six-factor model to alternatives in which each construct correlation was sequentially constrained to unity revealed that the full model was superior in all cases. Chi-square difference tests were also used to assess the discriminant validity of the second-

[^3]TABLE 1
First-Order Factor Loadings of Fandim Items and Construct Reliabilities

| When watching a competition in this sport, I: |  | Study 2 ( $N=624$ ) |  |  |  | Study 3 ( $N=263$ ) |  |  |  | Study $4(N=372)$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $\lambda x^{2}$ | $C R^{\text {b }}$ | $\mathrm{AL}^{\text {c }}$ | $\mathrm{AV}^{\text {d }}$ | $\lambda x$ | CR | AL | AV | $\lambda x$ | CR | AL | AV |
| AES 1 | admire the artistry displayed in the sport ${ }^{\text {e }}$ | . 84 | . 92 | . 92 | . 73 | . 84 | . 93 | . 90 | . 70 | . 88 | . 94 | . 94 | . 79 |
| AES 2 | admire the beauty of the sport | . 88 |  |  |  | . 87 |  |  |  | . 83 |  |  |  |
| AES 3 | am moved by the gracefulness of the sport | . 86 |  |  |  | . 85 |  |  |  | . 92 |  |  |  |
| AES 4 | focus on the elegance of the sport | . 83 |  |  |  | . 78 |  |  |  | . 91 |  |  |  |
| EVA 1 | critically evaluate the performance of the teams/athletes | . 84 | . 90 | . 90 | 69 | . 88 | . 90 | 90 | . 68 | . 93 | . 92 | . 92 | . 74 |
| EVA 2 | analyze the performance of the teams and/or athletes | 86 |  |  |  | . 82 |  |  |  | . 79 |  |  |  |
| EVA 3 | make value judgments about the quality of a team/athlete's performance | . 82 |  |  |  | . 79 |  |  |  | . 77 |  |  |  |
| EVA 4 | critique the quality of a team/athlete's performance | . 80 |  |  |  | . 81 |  |  |  | . 93 |  |  |  |
| FAN 1 | imagine that I am one of the athletes | 90 | . 94 | . 94 | . 81 | . 89 | . 94 | . 94 | . 79 | . 86 | . 92 | . 95 | . 82 |
| FAN 2 | fantasize that I am participating in the action | 91 |  |  |  | . 87 |  |  |  | . 90 |  |  |  |
| FAN 3 | daydream that I am competing in the contest | . 92 |  |  |  | . 92 |  |  |  | . 93 |  |  |  |
| FAN 4 | envision myself as one of the competitors | . 86 |  |  |  | . 88 |  |  |  | . 94 |  |  |  |
| FLO 1 | get so "lost" in the action that time seems to be altered | . 82 | . 92 | . 92 | . 75 | . 81 | . 93 | . 93 | . 76 | . 67 | . 92 | . 91 | . 74 |

TABLE 1
(Continued)

| When watching a competition in this sport, I: |  | Study $2(N=624)$ |  |  |  | Study $3(N=263)$ |  |  |  | Study $4(N=372)$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $\lambda x^{2}$ | $\mathrm{CR}^{\text {b }}$ | $\mathrm{AL}^{\text {c }}$ | $\mathrm{AV}^{\text {d }}$ | $\lambda x$ | CR | AL | AV | $\lambda x$ | CR | AL | AV |
| FLO 2 | get so into the action that I lose touch with what is happening around me | . 91 |  |  |  | . 85 |  |  |  | . 86 |  |  |  |
| FLO 3 | feel as if time is standing still because I'm so focused on the action | . 89 |  |  |  | . 88 |  |  |  | . 93 |  |  |  |
| FLO 4 | am so "zoned into" the action that I lose sense of time | . 84 |  |  |  | . 95 |  |  |  | . 95 |  |  |  |
| PER 1 | pay attention to only those athletes who are most well known | . 89 | . 94 | . 94 | . 79 | . 83 | . 94 | . 94 | . 79 | . 71 | . 92 | . 91 | . 74 |
| PER 2 | focus only on those athletes who are most famous | . 93 |  |  |  | . 91 |  |  |  | . 88 |  |  |  |
| PER 3 | am attentive only to those athletes with the most notoriety | . 85 |  |  |  | . 89 |  |  |  | . 90 |  |  |  |
| PER 4 | concentrate on only those athletes who are most prominent | . 88 |  |  |  | . 93 |  |  |  | . 92 |  |  |  |
| PHY 1 | admire the physiques of the athletes while they are performing | . 84 | . 91 | . 91 | . 72 | . 86 | . 92 | . 92 | . 74 | . 86 | . 91 | . 91 | . 72 |
| PHY 2 | focus on the "sex appeal" of the athletes as they compete | . 86 |  |  |  | . 77 |  |  |  | . 74 |  |  |  |
| PHY 3 | admire the bodies of the athletes as they compete | . 89 |  |  |  | . 88 |  |  |  | . 87 |  |  |  |
| PHY 4 | am captivated by the appeal of the athletes' bodies as they perform | . 81 |  |  |  | . 91 |  |  |  | . 92 |  |  |  |

Key: AES $=$ Aesthetics, EVA $=$ Evaluation, FAN $=$ Fantasy, FLO $=$ Flow, PER $=$ Personalities, PHY $=$ Physical Attractiveness
${ }^{\text {a }}$ Completely standardized parameter.
${ }^{\mathrm{b}}$ Construct reliability computed as $(\Sigma \lambda)^{2} /\left[(\Sigma \lambda)^{2}+\Sigma \operatorname{var}(\varepsilon)\right]$.
${ }^{\text {c }}$ Coefficient alpha.
${ }^{\text {d }}$ Average variance extracted, which is the proportion of variance in the construct that is not attributable to measurement error.
"All items measured on a 10 -point scale ranging from $1=$ not at all like me to $10=$ very much like me. All item loadings on their respective factors were significant at the .001 level.

TABLE 2
Summary Table of All Sex Differences on Subscales by Sports Group Attaining a Significance Level of .05 or Less ( $N=1259$ )

| Sports Group | Two-Way Interaction ${ }^{\text {a }}$ <br> [F, (df), p-value] | SIMPLE EFFECT TESTS FOR SEX ON SUMMED SUBSCALES WITHIN EACH SPORTS GROUP [ F, (df), p-value] |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | AES | EVA | FAN | FLO | PER | PHY |
| Aesthetic Sports Performed | 14.22 | $\mathrm{F}>\mathrm{M}^{\text {b }}$ | F $>\mathrm{M}$ | $\mathrm{F}>\mathrm{M}$ |  |  | $\mathrm{M}>\mathrm{F}$ |
| by Women | $(5,289)$ | $10.25^{c}(1,293)$ | 4.19 (1, 293) | $8.50(1,293)$ |  |  | 23.03 (1, 293) |
| ( $n=295$ ) | $<.001$ | $<.001$ | $<.04$ | <. 01 |  |  | $<.001$ |
| Aesthetic Sports Performed | 12.26 | $\mathrm{F}>\mathrm{M}$ | $\mathrm{F}>\mathrm{M}$ |  |  |  | $\mathrm{F}>\mathrm{M}$ |
| by Men | $(5,277)$ | 16.76 (1, 281) | 5.15 (1, 281) |  |  |  | 50.04 (1, 281) |
| ( $n=283$ ) | $<.001$ | $<.001$ | $<.03$ |  |  |  | $<.001$ |
| Purposive Sports Performed | 10.26 |  |  |  | $\mathrm{M}>\mathrm{F}$ | $\mathrm{M}>\mathrm{F}$ | $\mathrm{M}>\mathrm{F}$ |
| by Women | $(5,303)$ |  |  |  | 8.43 (1, 307) | 7.62 (1, 307) | 51.78 (1, 307) |
| ( $n=309$ ) | $<.001$ |  |  |  | $<.01$ | $<.01$ | $<.001$ |
| Purposive Sports Performed | 51.41 | $\mathrm{M}>\mathrm{F}$ | $\mathrm{M}>\mathrm{F}$ | $\mathrm{M}>\mathrm{F}$ | $\mathrm{M}>\mathrm{F}$ |  | $\mathrm{F}>\mathrm{M}$ |
| by Men | $(5,366)$ | 28.50 (1, 370) | 56.94 (1, 370) | 46.01 (1, 370) | 23.43 (1, 370) |  | 91.36 (1, 370) |
| ( $n=372$ ) | $<.001$ | $<.001$ | $<.001$ | $<.001$ | <.001 |  | $<.001$ |

Key: $\mathrm{AES}=$ Aesthetics, EVA $=$ Evaluation, FAN $=$ Fantasy, $\mathrm{FLO}=$ Flow, $\mathrm{PER}=$ Personalities, PHY $=$ Physical Attractiveness
${ }^{\text {a }}$ Multivariate Wilk's lambda for two-way interactions (SUBSCALE $\times$ SEX) from repeated measures analysis within each sports group.
${ }^{6}$ Direction of differences attaining an F term resulting in a $p$-value of .05 or less; $\mathrm{M}=$ Males, $\mathrm{F}=$ Females.
${ }^{c}$ Univariate $F$ value indicating sex differences on the mean dimension score within each sports group.
order factors. For each sample, the chi-square value yielded by the secondorder analysis was compared to one from a model in which the correlation between the higher-order factors was constrained to unity. Based on chisquare difference tests, the constrained model provided a significantly worse fit to the data than did the unconstrained model in each sample (all $p$ 's $<$ .001). In addition, none of the higher-order factor covariances were within two standard errors of one. In sum, discriminant validity was evidenced for both the first- and second-order factors in each sample.

## Testing for Factorial Invariance

Although the data from each sample appear to be best represented by the higher-order model, questions remain as to whether differences in fit exist across samples or by sex. Thus, two sets of invariance analyses were conducted (see Steenkamp and Baumgartner, 1998).

Assessing FANDIM's measurement invariance across samples. ${ }^{5}$ The best fitting two-factor higher-order model from each sample (M8 for Study 2 data, M7 for all other samples) was compared across groups. The results of invariance tests indicated the same pattern of fixed and free elements across the three samples (i.e., configural invariance), no differences in factor loadings across groups (i.e., metric invariance), and the intercepts of the measured variables were equivalent across samples (i.e., scalar invariance). ${ }^{6}$ The presence of configural, metric and scalar invariance is necessary to compare latent mean scores. A test of full error invariance was rejected. It should be noted that error invariance is not required to compare latent mean scores. Subsequent tests of partial error variance invariance were conducted that involved sequentially relaxing three non-invariant error variances in the Study 3 data and four non-invariant variances in the Study 4 data. The final model of partial error invariance revealed no group differences.

Invariance tests of first-order factor loadings on their respective secondorder factors revealed two non-invariant loadings in the Study 3 sample. Relaxing the loadings of FAN and FLO on Autotelism in this sample resulted in an equivalent fit across groups. A test of second-order factor covariance invariance indicated that the correlation between second-order constructs across samples was non-invariant based on the CAIC value. Finally, the initial test of latent factor mean invariance was rejected. However, a test of partial latent factor mean invariance was supported after relaxing the invariance constraint for Appreciation in the Study 3 sample.

[^4]Assessing FANDIM's measurement invariance by sex. Using M7 as the baseline comparator, invariance tests by respondent sex revealed invariance in all but one case. The latent factor mean for Autotelism was greater for men than it was for women.

Summary of tests of factorial invariance. On the basis of tests for configural, metric and scalar invariance, the results suggest that the fit of the twofactor model was equivalent across the three samples. Similarly, no meaningful differences were found based on sex. Given model equivalence, the data from the three studies were combined and summed scores were created for each of the six subscales (FAN, FLO, EVA, PER, PHY, AES) and two higherorder factors (Autotelism, Appreciation). These values were then used in subsequent analyses.

## Dimension Differences Attributable to Sports Group and Sex

Evidence has emerged in recent years indicating that men and women differ in their behaviors toward and enjoyment of spectator sports. Gantz and Wenner (1991) reported that compared to women, men are more likely to attend sporting events, read about sports, watch sports-related news stories on television, and also possess greater knowledge about sports. Bryant and Zillmann (1983) found that men derived greater enjoyment than women from watching the rough plays exhibited in professional football. Differences were also found by Sargent, Zillmann, and Weaver (1998) who reported that women were more apt to enjoy elegant or stylistic sports (e.g., gymnastics, figure skating) whereas men found greater pleasure in watching combative sports (e.g., football, hockey). Men also derived greater enjoyment from sports featuring mechanical elements such as golf and mountain biking.

The preceding suggests that sex is an important individual difference variable that at least partially explains how a spectator experiences a sporting event. It therefore follows that just as Sargent et al.'s (1998) study reported sex differences in enjoyment based on specific properties embodied in a sport, so too should differences exist along the dimensions used by spectators to consume a sporting event. Thus, a $2 \times 4 \times 6$ overall mixed-design ANOVA was conducted in which sex and sports group were between-subjects variables and the six FANDIM subscales were treated as levels of a within-subjects factor. Cell sizes ranged from 71 to 97 . The four sports groups were based on whether the sport was performed by men or women and the classification of the sport as either aesthetic or purposive. The analysis yielded significant main effects for the subscales factor (Wilk's $F=359.88$; df $=5,1247 ; p<$ .001) and for the between-subjects measure of Sports Groups ( $F=10.15$; df $=3,1251 ; p<.001$ ). No main effect was found for $\operatorname{Sex}(p>.35)$. In addition, two-way interactions were found for Sports Groups $\times$ Dimension (Wilk's $F$ $=27.62 ; \mathrm{df}=15,3443 ; p<.001$ ), Sex $\times$ Dimension (Wilk's $F=3.61 ; \mathrm{df}=$ 5, 1247; $p=.001$ ), and Sports Groups $\times \operatorname{Sex}(F=11.96 ; \mathrm{df}=3,1251 ; p<$ .001). However, each of these lower-order effects was qualified by a significant

Sex $\times$ Sports Groups $\times$ Dimension interaction (Wilk's $F=24.87$; $\mathrm{df}=15$, 3443; $p<.001$ ).

Follow-up tests were conducted to determine whether sex differences existed in the dimension scores within each sport group. The left hand column of Table 2 features the repeated measures two-way interactions involving Sex $\times$ Dimension conducted within each of the four sports groups. All interactions exceeded the .001 significance level. The data featured in the remaining columns are the results of the simple-effects tests examining sex differences for each dimension. The direction of differences is shown for those tests yielding $F$ scores with $p<.05$. Of the 24 comparisons, 15 ( $63 \%$ ) yielded significant $p$-values. Men's scores were significantly higher than women's on eight of these comparisons. Men scored higher than women on four (i.e., AES, EVA, FAN, FLO) of the six dimensions related to purposive sports performed by men and on three dimensions of purposive sports performed by women (i.e., FLO, PER, PHY). Significant sex differences were found for PHY in all four sports groups. Higher PHY scores were recorded for men than women when consuming women's sports (both purposive and aesthetic) and the converse was true of women consuming men's sports. Women scored higher than men on the AES and EVA dimensions for men's and women's aesthetic sports; as well as on the FAN dimension for aesthetic sports performed by women.

## Likelihood of Watching and the FANDIM Dimensions

The next analysis considers the extent to which dimension scores are correlated with likelihood of watching the particular sport for which the FANDIM was completed. Intragroup correlations were calculated between each dimension and likelihood of watching within each sports group. The intragroup correlations were then transformed for comparison using Fisher's $r$ to-z procedure (Cohen \& Cohen, 1983). The use of Fisher's $r$-to-z is appropriate because the sport-type groups were independent from one another. Intragroup correlation comparisons revealed one significant difference based on respondent sex. The transformed correlation between EVA and likelihood of watching (i.e., EVA/watch correlation) for those completing the scale in regard to an aesthetic sport performed by women was greater for male respondents ( $r$ to- $z=.50, p<.001$ ) than it was for female respondents ( $r$ to- $z=.25, p<.01 ; z=2.11, p<.05$ ). This suggests that male interest in watching an aesthetic sport performed by women is more highly linked to evaluating the action than is female interest in watching this type of sport.

Given that only one difference was found, transformed correlations aggregated across sex were calculated and comparisons made between (a) male and female purposive sports groups and (b) male and female aesthetic sports groups. The analyses consider whether differences exist within a sports group according to the sex of the performer. Only one difference was found. The PHY/watch correlation was significantly greater ( $r$-to- $-=.32, p<.001$ ) for those in the aesthetic-sports-performed-by-men group than it was for those
in the aesthetic-sport-performed-by-women group ( $r$-to- $z=.13, p<.03 ; z=$ $-2.26, p<.03$ ). Thus, regardless of respondent sex, interest in watching an aesthetic sport performed by men was more positively related to physical attractiveness than it was to watching an aesthetic sport performed by women.

Next, the data were aggregated again, this time by respondent sex and performer sex. For aesthetic sports, the results revealed significant correlations between each FANDIM dimension and likelihood of watching ( $r$-to- $z$ 's: FAN $=.41, \mathrm{FLO}=.51, \mathrm{EVA}=.41, \mathrm{AES}=.43, \mathrm{PER}=.18, \mathrm{PHY}=.24$; all $p$ 's $<.001$ ). Four of the six correlations for purposive sports were significant ( $r$-to-z's: FAN $=.30, \mathrm{FLO}=.47, \mathrm{EVA}=.46$, AES $=.36$; all $p$ 's $<.001$ ) and two were not ( $r$-to-z's: PER $=.02, \mathrm{PHY}=.03$; each $p>.50$ ). Comparing each dimension/watch correlation across sport type (i.e., aesthetic vs. purposive) revealed two differences. The PER/watch correlation for those completing the scale in regard to an aesthetic sport was significantly greater than was that correlation for those in the purposive-sports group, $z=2.81, p<.01$. This suggests that, in contrast to purposive sports, increased attending to the personalities performing an aesthetic sport is positively related to interest in watching.

The second significant difference between aesthetic and purposive sport types was found in the PHY/watch correlation ( $z=3.68, p<.001$ ). However, due to the significant sex-of-performer difference in this correlation for aesthetic sports found earlier, two separate analyses were conducted. First, the PHY /watch correlation for those in the purposive-sports group was compared to the same correlation in the aesthetic-sport-performed-by-men group. Second, the same comparison was made for those in the aesthetic-sport-performed-by-women group. Regarding the former, the correlation for those in the aesthetic-sport-performed-by-men group ( $r$-to- $z=.32, p<.001$ ) was significantly greater than the corresponding correlation for those in the aggregated purposive sport-type group ( $r$ to- $z=.03 ; z=4.09, p<.001$ ). The second comparison revealed a significant difference indicating that those in the aesthetic-sport-performed-by-women group had a greater correlation coefficient ( $r$ to $-z=.13$ ) than those in the aggregated purposive sport-type group ( $r$ to- $z=.03 ; z=1.43, p<.06$ ). Thus, attending to the physical attractiveness of the athletes was more highly related to likelihood of watching an aesthetic sport performed by either men or women than it was to watching a purposive sport.

## Study 4: Tests of Concurrent Validity

Data collected as part of Study 4 also allowed for comparisons to be made between the FANDIM'S dimensions and a variety of other constructs. Respondents completed the five-item Experiential Response (ER) and threeitem Imaginal Response (IR) scales (Lacher \& Mizerski, 1994). Item wording was adapted slightly to fit the current context. Respondents were asked to complete each scale in reference to the same sport selected when completing

TABLE 3
Pearson Correlations of Summed Fandim Dimensions with Other Measures

|  | Hedonic <br> Response $^{\mathrm{a}}$ <br> $(N=372)$ | Involvement ${ }^{\mathrm{b}}$ <br> $(N=266)$ | Hedonic <br> Motives $^{\mathrm{d}}$ <br> $(N=266)$ | Economic <br> Motives $^{\mathrm{e}}$ <br> $(N=266)$ | Aesthetic <br> Motives $^{\mathrm{s}}$ <br> $(N=266)$ |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Fantasy | .51 | $-.44^{\mathrm{c}}$ | .48 | .28 | .23 |
| Flow | .61 | -.52 | .61 | .23 | $.15(.012)$ |
| Evaluation | .50 | -.45 | .42 | $.09(.14)$ | $.20(.001)$ |
| Personalities | .06 | .00 | $.11(.07)$ | .28 | $-.03(.66)$ |
| Physical Attractiveness | $.16(.002)$ | $-.10(.09)$ | $.08(.19)$ | $.14(.03)$ | $.19(.003)$ |
| Aesthetics | .36 | -.35 | .30 | $.07(.29)$ | .71 |
| AUTOTELISM | .68 | -.58 | .62 | .25 | .24 |
| APPRECIATION | .27 | -.22 | .23 | .22 | .54 |
| $t$ value (df) ${ }^{\mathrm{g}}$ | $11.00(369)$ | $7.48(263)$ | $8.43(263)$ | $.40(263)$ | $-3.22(263)$ |
| $p<$ | .001 | .001 | .001 | NS | .001 |

Note: All p-values are at the .001 level of significance unless otherwise noted in parentheses.
${ }^{2}$ Hedonic Response is a summed eight-item measure consisting of the Experiential Response and Imaginal Response scales from Lacher and Mizerski (1994).
${ }^{\mathrm{b}}$ Involvement is a summed 11 -item measure from Bloch (1981). Items are worded such that lower values indicate greater involvement.
${ }^{\circ}$ Negative correlations indicate a positive relationship between involvement and the corresponding dimension.
${ }^{\mathrm{d}}$ Hedonic Motives is a summed nine-item scale consisting of items from the following of Wann's (1995) SFMS subscales: escape, eustress, self-esteem.
${ }^{\text {e}}$ Economic Motives is a summed three-item subscale from Wann (1995).
${ }^{\mathrm{f}}$ Aesthetic Motives is a summed three-item subscale from Wann (1995).
${ }^{\text {g }}$ The t-test compares correlation coefficients. Thus, the first test compares the correlation between Hedonic Response and the 12 -item Autotelism measure (summed items from the Fantasy, Flow and Evaluation subscales) to the correlation between Hedonic Response and the 12 -item Appreciation measure (summed items from the Personalities, Physical Attractiveness and Aesthetics subscales).
the FANDIM scale. The ER scale considers the extent to which a person becomes absorbed or carried off in the consumption experience and the IR scale is designed to capture the fantasy aspect of hedonic consumption. Thus, both scales should be more highly related to Autotelism and its constituent dimensions (i.e., FAN, FLO, EVA) than to Appreciation and its dimensions.

Approximately one week after data collection, all Study 4 respondents were sent a follow-up email that included a link to two additional scales. The sport for which the FANDIM scale was originally completed was also referenced in the email and respondents were asked to complete each scale in regard to this sport. A separate question included on the survey asked respondents to indicate the referenced sport. All respondents indicated the correct sport. In contrast to the original Study 4 data collection, respondents
received no extra credit for participation. Complete data were returned from 266 individuals for a response rate of $72 \%$. The presentation of the scales to respondents was varied to reduce ordering effects. The two scales included Bloch's (1981) 17-item Involvement with a Product Class (IPC) scale and Wann's (1995) 23 -item Sport Fan Motivation Scale (SFMS). As before, respondents were asked to complete the scales in reference to the sport selected in the Study 4 collection. The IPC was used to evaluate the consumer's long-term interest in the selected sport. Originally developed to examine the centrality of an automobile to a consumer's values, needs and self-concept, the wording of the IPC was adapted to fit a spectator sport context. Bloch reported that the IPC was represented by six underlying factors: enjoyment and usage of cars; readiness to talk to others about cars; interest in car racing activities; self-expression through one's car; attachment to one's car; interest in cars. The SFMS includes eight subscales (eustress, self-esteem, escape, entertainment, economic, aesthetic, group affiliation, and family needs) thought to be motives for sport fandom. All but the two-item Family Needs subscale was included in the survey.

Lacher and Mizerski (1994): Hedonic Response. The eight items comprising ER and IR were submitted to an oblique rotation factor analysis. An examination of the eigenvalues and the scree plot revealed the existence of a single dominant factor which accounted for $62.9 \%$ of the explained variance (Cronbach's $\alpha=.92$ ). The eight items comprising the factor, now referred to as Hedonic Response (HR), were summed and pairwise correlations were made with the FANDIM scales. The results are shown in Table 3. Consistent with expectations, the correlations of HR and each of the Autotelism dimensions was significant; whereas correlations with the Appreciation dimensions were substantially lower. Table 3 also features the correlation between HR and each of the summed higher-order dimensions. Not surprisingly, the results of a $t$ test comparing dependent correlation coefficients (see Cohen \& Cohen, 1983, pp. 56-57) revealed that HR was significantly more correlated with Autotelism than with Appreciation.

Bloch (1981): Involvement. An oblique factor analysis of the 17 -item scale yielded a scree plot indicating the presence of two strong factors. The first factor, which explained $44 \%$ of the variance, contained 11 items related to enjoyment, readiness, attachment, and self-expression. One of the 11 items cross-loaded on an uninterpretable second factor which also contained each of the reverse-coded items. Hence, the 10 items loading cleanly on the first factor were summed to represent Involvement (INV, $\alpha=.92$ ). Due to scale wording, the negative coefficients shown in Table 7 indicate that higher scores on Autotelism and each of its dimensions (FAN, FLO, EVA) are related to increased levels of INV. As with HR, INV's correlation with the summed Autotelism measure was significantly greater than its correlation with Appreciation.

Wann (1995): SFMS. Although the results of an oblique factor analysis indicated four factors with eigenvalues greater than 1.0, the scree plot revealed only three strong factors which accounted for $64 \%$ of the variance.

Items comprising the escape, eustress and self-esteem subscales loaded on the first factor ( $39.6 \%$ of the variance, $\alpha=.91$ ). The escape motive is concerned with fans' desire to watch a sporting event because it serves as a diversion from the mundane routines associated with everyday life. Eustress refers to the enjoyable stress resulting from the arousal associated with an upcoming sporting event. The self-esteem motive addresses the vicarious achievement derived by fans as a result of a team's performance. Collectively, the three subscales reflect fans' hedonic motives for watching a sport and the factor was named accordingly (i.e., Hedonic Motives, HM). The second factor included the three items comprising the Economic Motives (EM) subscale ( $13.4 \%$ of the variance, $\alpha=.87$ ) which posits that fans are motivated by the potential economic gains garnered through betting on the outcomes of sporting events. The final interpretable factor consisted of the three items from the Aesthetic Motives (AM) subscale ( $10.9 \%$ of the variance, $\alpha=.91$ ). This subscale is intended to assess the extent to which fans are motivated to watch because of the artistry of the sport and the creativity of its athletes. The remaining items from the group affiliation and entertainment subscales all loaded on a single uninterpretable factor that did not make a unique contribution to the scree plot.

As indicated in Table 3, the HM measure is positively related to the three dimensions underlying Autotelism. It thus follows that HM is more highly correlated with the higher-order Autotelism measure than with Appreciation. The EM scale was significantly related to FAN, FLO and PER; but no differences were found in the correlations between EM and either of the higherorder measures. Wann's AM subscale was highly correlated with FANDIM's AES subscale. Interestingly, this measure was also positively related to EVA and PHY.

In sum, Autotelism and its constituent dimensions were positively related to the following scales: Wann's (1995) Hedonic Motives, Bloch's (1981) Involvement, and Lacher and Mizerski's (1994) Hedonic Response. These results are consistent with the way in which autotelism was conceptualized in the FANDIM scale. Autotelism refers to consumption for its own sake and is focused on the consumer's interaction with the performance itself. Not surprisingly, Wann's (1995) Aesthetic Motives factor was highly correlated with FANDIM's Aesthetic dimension. In addition, AES was also positively related to each of the other measures except Economic Motives. Although the correlations were less than those between these measures and the Autotelism dimensions, the positive relationships indicate that an appreciation for the beauty of the sport also taps the hedonic value of sporting event consumption.

## Study 5: Test-Retest

Test-retest coefficients were calculated for the summed scale and for each of the dimensions. The Pearson correlation between the two data collections for the summed scale was .91 . The mean correlation across the six
subscales calculated individually was .77 with a range of .71 (AES, FAN, PHY) to .87 (FLO). Thus, the FANDIM and its constituent dimensions demonstrated adequate test-retest reliability.

## General Discussion

The current article describes the development of a scale designed to assess the underlying dimensions along which sporting event performances are consumed. The best fitting empirically supported model was one represented by two higher-order factors, each comprised of three distinct firstorder factors. The three first-order factors contributing to Autotelism include Fantasy, Flow and Evaluation; whereas Personalities, Physical Attractiveness and Aesthetics contribute to the Appreciation factor. The FANDIM instrument was found to possess acceptable internal consistency, test-retest reliability and discriminant validity. In addition, no substantive differences in factor structure were observed between the three samples or on the basis of sex. Tests of concurrent validity conducted as part of Study 4 indicated that a number of the summed higher- and lower-order dimensions comprising FANDIM were correlated with other measures tapping similar theoretical domains. Autotelism and its associated first-order factors were positively related to two different measures of hedonic consumption, as well as overall involvement with the sport for which the FANDIM was completed. The correlation between each of these scales and the summed Autotelism measure was also significantly greater than that of the corresponding correlations with Appreciation. Wann's (1995) Aesthetic Motives measure for general interest in spectator sports was correlated with the more sport specific Aesthetic and Personalities subscales of the FANDIM. Moreover, Wann's Aesthetic Motives measure was significantly more positively correlated with Appreciation than it was with Autotelism.

Analyses using data aggregated across three of the five studies reported in this research revealed that female and male consumption of different types of sports varied along FANDIM's dimensions. For purposive sports performed by men, male-compared to female-respondents were more likely to (a) evaluate the performance of the athletes competing in the sport, (b) fantasize that they were part of the action, (c) achieve a sense of flow during competitions, and (d) appreciate the beauty of these sports. Women, on the other hand, were more likely to appreciate the beauty of aesthetic sports performed by either men or women and evaluate the performances of athletes in those sports than were their male counterparts. Women also fantasized that they were more a part of the action when consuming women's aesthetic sports than did male respondents. Also, not surprisingly, greater appreciation for the physical attractiveness of the athletes was found for females when consuming men's sports (purposive and aesthetic), whereas the converse existed for males consuming women's sports. However, as suggested by a reviewer, future research should examine the extent to which sex differences might be mitigated by other variables such as team identification.

A separate set of analyses also using the aggregated data revealed significant relationships between various subscales of the FANDIM and respondent likelihood of watching a particular type of sport for which the scale was completed. In the case of aesthetic sports, regardless of performer sex, each FANDIM subscale was positively related to likelihood of watching. For purposive sports, greater interest in watching was positively related to FAN, FLO, EVA, and AES. Interestingly, significant likelihood-of-watching differences between sports groups were found for correlations involving PER and PHY. In contrast to the correlations found for purposive sports, interest in watching aesthetic sports was significantly more correlated with respondent appreciation of the athletes' personalities and physical attractiveness.

In sum, based on aggregated data, the results suggest that the consumption of sporting event performance as measured along the FANDIM's dimensions does indeed vary by respondent sex and the sports groupings considered here. The FANDIM's dimensions are also related to interest in watching (i.e., consuming). Moreover, the relations between dimensions and interest in watching are also differentially affected by the sports groupings used in this study.

## Managerial Implications

The patterns of sports consumption described by the FANDIM scale should have relevance to a number of marketplace activities. The scale provides practitioners with a reliable tool that can be easily used to collect data about how fans consume sporting events. The information can then be used to create messages aimed at specific target markets. For example, based on the interest-in-watching correlations with PHY and PER across sports groups found in this research, marketers might want to place greater emphasis on the personal characteristics of individual athletes than their athletic prowess when promoting aesthetic sports. This could be achieved by creating vignettes that introduce viewers to a particular athlete. These vignettes could then be used in promoting the broadcast. At a more macro level, the various FANDIM dimensions might also be used as a basis for segmenting a market. Although not explicitly tested here, it is possible that different sports yield different consumption patterns across the six FANDIM dimensions.

## Limitations

The current work is limited by the fact that consumption was considered in retrospect rather than as a sporting event actually unfolds. Based on previous research indicating that experiential events are more accurately recalled when respondents are asked to focus on specific instances (see Conway, 1992, 1996), it was hoped that retrospective accuracy would be enhanced by asking respondents to complete the scale in regard to a specific sport rather than asking about their level of fanship for sports in general. Another concern was that asking people about how they are consuming an event in
real time might change the nature of the experience. It was also felt that a retrospective scale would be of greater value to practitioners conducting research in a field setting. Nevertheless, the retrospective nature of the scale may be problematic, especially for dimensions such as Flow. Future work should investigate whether dimension differences exist for people completing the scale immediately following or soon after an event compared to those completing it after a more extended period of time. A test-retest under these conditions would be quite informative as to the retrospective approach used here.

The decision to emphasize six dimensions of sporting event consumption was based on the extant literature and the need to balance the scale's contribution with parsimony. However, other dimensions were considered for inclusion. One such dimension was socialization. Fans often consume a sporting event in the company of others and rely on others to help interpret the action on the field (see Holt, 1995). Exclusion of this dimension was based primarily on the fact that it involves interpersonal relations. The FANDIM was designed to focus on the consumer's own consumption cognitions. Also not included was a self-esteem or vicarious achievement dimension, which reflects basking-in-reflected-glory (BIRG; see Madrigal, 1995). Cialdini's (1985) original conceptualization of BIRG relied on an individual's strategic decision to form a bond with a successful other (i.e., sports team) in order to enhance his or her own self esteem in the eyes of another. As such, vicarious achievement relies on a team's successful outcomes. The focus of the FANDIM is on the consumption experience, not the motivations antecedent to or the judgments consequent to consumption.

Another shortcoming of the research is the use of sports groups rather than the individual sports comprising these groups. In order to ensure adequate group sizes, sports of a similar type (i.e., aesthetic vs. purposive) were grouped by performer sex. This allowed for more robust analyses but may have sacrificed nuanced differences existing across different sports within a group. Future research may want to consider the scale more narrowly by specific sports.

## Conclusion

Conceptual models detailing experiential consumption (Hirschman \&c Holbrook 1982) and the consumption of performance (Barthes, 1972; Deighton, 1992) served as the inspiration for developing the FANDIM scale. Although previous work addresses specific motives for watching sports generally (see Gantz \& Wenner, 1991; Wann, 1995), only Holt's (1995) interpretive work offers a typology detailing how people actually consume a sporting event performance. The FANDIM scale was designed to capture the spirit and richness of these conceptual models in a way that would have practical implications for sports marketers.

The FANDIM offers one way of assessing how people consume a specific leisure experience. Although designed for elucidating sporting event con-
sumption, it would be interesting to examine the extent to which the scale's dimensions are generalizable to other types of performances such as stage plays or music concerts. Obviously, the wording of scale items would need to be adapted to do this. Nevertheless, the conceptual basis upon which the scale was constructed suggests that many of the dimensions may be generalizable. Understanding the myriad of ways that leisure experiences are consumed and examining how people differ on those dimensions represents a fascinating area for future research.

## References

Barthes, R. (1972). Mythologies. New York: Hill and Wang.
Belk, R. W., \& Costa, J. A. (1998). The mountain man myth: A contemporary consuming fantasy. Journal of Consumer Research, 25, 218-240.
Best, D. (1978). The aesthetic in sport. In D. Best (Ed.), Philosophy and human movement (pp. 99122). London: Allen and Unwin.

Bloch, P. H. (1981). An exploration into the scaling of consumers' involvement with a product class. In E. B. Monroe (Ed.), Advances in consumer research, vol. 8 (pp. 61-65). Provo, UT: The Association for Consumer Research.
Boxill, J. M. (1985). Beauty, sport, and gender. Journal of the Philosophy of Sport, 11, 36-47.
Bozdogan, H. (1987). Model selection and Akaike's information criteria (AIC): The general theory and its analytical extensions. Psychometrika, 52, 345-370.
Bryant, J., \& Zillmann, D. (1983). Sports violence and the media. In J. H. Goldstein (Ed.), Sporis violence (pp. 195-211). New York: Springer-Verlag.
Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., \& Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. Journal of Personality and Social Psychology, 34, 366-375.
Cohen, J., \& Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral sciences ( $2^{\text {nd }}$ ed.). New York: Wiley.
Conway, M. A. (1992). A structural model of autobiographical memory. In M. A. Conway, D. C. Rubin, H. Spinnler, \& W. A. Wagenaar (Eds.), Theoretical perspectives on autobiographical memory (pp. 167-193). Dordrecht, Netherlands: Kluwer.
Conway, M. A. (1996). Autobiographical knowledge and autobiographical memories. In D. C. Rubin (Ed.), Remembering our past: Studies in autobiographical memory (pp. 67-93). New York: Cambridge University Press.
Csikszentmihalyi, M. (1975). Beyond boredom and anxiety. San Francisco: Jossey-Bass.
Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. New York: Harper and Row.
Deighton, J. (1992). The consumption of performance. Journal of Consumer Research, 19, 362-372.
DeVellis, R. F. (1991). Scale development: Theory and applications. Newbury Park, CA: Sage.
Dietz-Uhler, B., Harrick, E. A., End, C., \& Jacquemotte, L. (2000). Sex differences in sport fan behavior and reasons for being a sport fan. Journal of Sport Behavior, 23, 219-231.
Duncan, M. C., \& Brummett, B. (1989). Types and sources of spectating pleasure in televised sports. Sociology of Sport Journal, 6, 195-211.
Fornell, C., \& Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18, 39-50.
Ganahl, J. (2004, August 22). Best reason to watch the Olympics-those Speedos. San Francisco Chronicle, pg. F7.
Gantz, W., \& Wenner, L. A. (1991). Men, women, and sports: Audience experiences and effects. Journal of Broadcasting and Electronic Media, 35, 233-243.
Gutmann, A. (1986). Sports spectators. New York: Columbia University.

Guttmann, A. (1996). The erotic in sports. New York: Columbia University.
Hansen, H., \& Gauthier, R. (1989). Factors affecting attendance at professional sport events. Journal of Sport Management, 3(1), 15-32.
Havitz, M. E., \& Dimanche, F. (1999). Leisure involvement revisited: Drive properties and paradoxes. Journal of Leisure Research, 31, 122-149.
Hirschman, E. C., \& Holbrook, M. B. (1982). Hedonic consumption: Emerging concepts, methods and propositions. Journal of Marketing, 46, 92-101.
Hofacre, S. (1994). The women's audience in professional indoor soccer. Sport Marketing Quarterly, 3(2), 25-27.
Holbrook, M. B., \& Zirlin, R. (1985). Artistic creation, artworks, and aesthetic appreciation. In R. Belk (Ed.), Advances in nonprofit marketing (pp. 1-54). Greenwich, CT: JAI.

Holt, D. B. (1995). How consumers consume: A typology of consumption practices. Journal of Consumer Research, 22, 1-16.
$\mathrm{Hu}, \mathrm{L} .$, \& Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6(1), 1-55.
Jackson, S. A., \& Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The flow state scale. Journal of Sport and Exercise Psychology 18(1), 17-35.
James, J. D., \& Ridinger, L. L. (2002). Female and male sport fans: A comparison of sport consumption motives. Journal of Sport Behavior, 25, 260-278.
Kahle, L. R., Kambara, K. M., \& Rose, G. (1996). A functional model of fan attendance motivations for college football. Sport Marketing Quarterly, 5(4), 51-60.
Lacher, K. T., \& Mizerski, M. (1994). An exploratory study of the responses and relationships involved in the evaluation of, and in the intention to purchase new rock music, Journal of Consumer Research, 21 (September), 366-380.
Laurent, G., \& Kapferer, J. N. (1985). Measuring consumer involvement profiles. Journal of Marketing Research, 22(1), 41-53.
Madrigal, R. (1995). Cognitive and affective determinants of fan satisfaction. Journal of Leisure Research, 27, 205-227.
Madrigal, R. (2003). Investigating an evolving leisure experience: Antecedents and consequences of spectator affect during a live sporting event. Journal of Leisure Research, 35, 23-48.
Madrigal, R. (2004). A review of team identification and its influence on consumers' responses toward corporate sponsors. In L. R. Kahle, C. Riley (Eds.), Sports marketing and the psychology of marketing communication (pp. 241-255). Mahwah, NJ: Lawrence Erlbaum Associates.
Morse, M. (1983). Sport on television: Replay and display. In E. A. Kaplan (Ed.), Regarding television: Critical approaches-an anthology (pp. 44-66). Maryland: University Publications of America in association with The American Film Institute.
Mulvey, L. (1975). Visual pleasure and narrative cinema. Screen, 16(3), 6-18.
Sargent, S. L., Zillmann, D., \& Weaver, J. B. III (1998). The gender gap in the enjoyment of televised sports. Journal of Sport and Social Issues, 22(1), 46-64.
Shields, R. (1992). Spaces for the subject of consumption. In R. Shields (Ed.), Lifestyle shopping: The subject of consumption (pp. 1-19). London: Routledge.
Silver, M. (1997, February 21). Beauty and the beach. Sports Illustrated, 216-231.
Sloan, L. R. (1989). The motives of sports fans. In J. H. Goldstein (Ed.), Sports, games and play: Social and psychological viewpoints (2 ${ }^{\text {nd }}$ ed., pp. 175-240). Hillsdale, NJ: Lawrence Erlbaum, 175-240.
Steenkamp, J-B. E. M., \& Baumgartner, H. (1998). Assessing measurement invariance in crossnational consumer research. Journal of Consumer Research, 25, 78-90.
Wann, D. L. (1995). Preliminary validation of the sport fan motivation scale. Journal of Sport and Social Issues, 19, 377-396.

Wann, D. L., Melnick, M. J., Russell, G. W., \& Pease, D. G. (2001). Sport fans: The psychology and social impact of spectators. New York: Routledge.
Wann, D. L., Schrader, M. P., \& Wilson, A. M. (1999). Sport fan motivation: Questionnaire validation, comparisons by sport, and relationship to athletic motivation. Journal of Sport Behavior, 22, 114139.
West, S. G., Finch, J. F., \& Curran, P. J. 1995. structural equation models with nonnormal variables: Problems and remedies. In Structural Equation Modeling: Concepts, Issues, and Applications. Ed. Rick H. Hoyle. Thousand Oaks, CA: Sage, 56-75.
Zillmann, D., Bryant, J., \& Sapolsky, B. S. (1989). Enjoyment from sports spectatorship. In J. H. Goldstein (Ed.), Sports, games, and play: Social and psychological viewpoints ( $2^{\text {nd }}$ ed., pp. 241278). Hillsdale, NJ: Lawrence Erlbaum.

Zillmann, D., \& Paulus, P. B. (1993). Spectators: Reactions to sports events and effects on athletic performance. In R. N. Singer, \& K. Tennant (Eds.), Handbook of research on sport psychology (pp. 600-619). New York: Macmillan.


[^0]:    Address correspondence to: Robert Madrigal, Associate Professor, Charles H. Lundquist College of Business, University of Oregon, 1208 University of Oregon, Eugene, OR, 97403-1208. Phone: (541) 346-5163. Fax: (541) 346-3341. Email: rmadriga@oregon.uoregon.edu

    Author note: The author gratefully acknowledges the help of Dennis Howard for his insight and encouragement in the development of the ideas presented in this paper.

[^1]:    ${ }^{1}$ The ordering of scale items was reversed for half of the subjects in each data set. The results of multiple $t$ tests comparing the mean level of responses for each construct by the ordering groups and a separate test examining the correlation of item ratings with the descriptors' sequence number on the scale revealed no evidence of systematic order bias in any of the studies.

[^2]:    ${ }^{2}$ An initial analysis of normality in all three samples using PRELIS 2.53 indicated non-normal distributions for a number of items. In such cases, West, Finch and Curran (1995) suggest that a Satorra-Bentler SCALED chi-square statistic be used. This adjustment was made for the data collected from each sample in Studies 2-4.

[^3]:    ${ }^{3}$ Summary of fit statistics for M7 for Studies 2, 3, 4, respectively: $\chi^{2}$ s $=589.90,310.11,469.41$; $\mathrm{dfs}=245$; RMSEAs $(90 \%$ confidence interval) $=.048$ (.043-.052), .032 (.019-.042), . 05 (.043.056 ); CFIs $=.98, .99, .98$. Fit statistics for M6 for each study, respectively: $\chi^{2} s=498.01,288.67$, $433.78 ; \mathrm{dfs}=237$; RMSEAs $(90 \% \mathrm{CI})=.042(.037-.047), .029(.014-.040), .047(.04 .054) ;$ SRMRs $=.035, .037, .041$; CFIs $=.98, .99, .98$.
    ${ }^{4}$ Allowing AES to cross load is justified for two reasons. First, no previous research has been conducted on these dimensions. Thus, the cross loading of AES may be an accurate representation of the phenomenon. Second, a case could be made that AES reflects on Autotelism because enjoying the beauty of the sport is an end unto itself.

[^4]:    ${ }^{5}$ A summary table of results of all invariance tests is available from the author upon request.
    ${ }^{6}$ A constrained model was considered non invariant if it yielded: (a) a RMSEA value outside the $90 \%$ CI of the unconstrained model and/or (b) a CAIC value greater than that of the unconstrained model.

