
Melanie L. Sartore
Texas A&M University

The purpose of this article is to examine the effects of treatment discrimination as manifested through performance evaluation bias. Research has demonstrated that demographic dissimilarity between raters and ratees can result in performance evaluation bias. As a result of such bias, work-related opportunities based on performance evaluation might be provided to individuals with demographic characteristics similar to those of the rater and not to those whose characteristics differ. The current model suggests that the denial of such opportunities is associated with the formation of self-limiting behavior and subsequent performance detriment. The current model also addresses the formation of a performance feedback loop that potentially leads to future poor performance. Moderators (i.e., task interdependence, time, social category identification, and common group identity) of the proposed relationships are also identified. Additionally, suggestions for future investigations are offered.

Discrimination represents one of the more pressing issues in organizations today, including those in sport and leisure. Indeed, several authors have noted the underrepresentation of women and ethnic minorities in coaching and upper-level management positions of sport organizations (e.g., Acosta & Carpenter, 2004; Cunningham, 2004a). Furthermore, Fink, Pastore, and Reimer (2001) suggested that within intercollegiate athletic departments, the “typical” employee is a White, Protestant, able-bodied, heterosexual male and that persons with characteristics different than these often faced difficult working environments. Lastly, Fink and Pastore (1999) noted that “perhaps nowhere is discrimination and oppression more evident than in Division IA intercollegiate athletics” (p. 311). Taken together, this literature suggests that (a) discrimination is prevalent within sport organizations and (b) persons of minority status attempting to enter and advance within sport organizations face numerous obstacles.

Greenhaus, Parasuraman, and Wormley (1990) identified two facets of discrimination: access discrimination, which excludes members of a particular group from entering an organization, and treatment discrimination, which occurs when members of a particular group within an organization receive fewer organizational resources,
rewards, and opportunities than are legitimately deserved. Indeed, both types of discrimination are applicable to multiple social groups as well as applicable across numerous contexts. For instance, discrimination is suggested to take place on the basis of demographic characteristics such as gender (Cunningham & Sagas, 2002; Ransom & Oaxaca, 2005; Reskin, 2000), ethnicity (Cunningham, 2004a; Reskin, 2000), and age (McVittie, McKinlay, & Widdicombe, 2003). Furthermore, several authors have argued that discrimination takes place within the various contexts, including higher education (Toutkoushian & Conley, 2005), business (Moyes, Williams, & Quigley, 2000), health care (Schigelone, 2003), and of particular relevance to this article, sport (Acosta & Carpenter, 2004; Cunningham, 2004a; Cunningham & Sagas, 2002; Fink et al., 2001; Sagas & Cunningham, 2004).

Extant research supports the notion that access discrimination exists within the sport context (Cunningham & Sagas, 2005; Lovett & Lowry, 1988; Stangl & Kane, 1991). Specifically, in an effort to understand the underrepresentation of women in the coaching profession, researchers have identified that women are discriminated within the hiring process through the practice of homologous reproduction (Lovett & Lowry, 1988; Stangl & Kane, 1991). Knoppers, Bedker-Meyer, Ewing, and Forrest (1989) demonstrated that women faced discrimination despite often possessing more human and social capital than their male counterparts, a finding that was later supported by Cunningham and Sagas’s (2002) investigation of intercollegiate basketball coaches. Similar findings of access discrimination within the sport context have also been documented among ethnic minority coaches (Anderson, 1993; Cunningham & Sagas, 2005).

In addition to the occurrence of access discrimination, there is also evidence of treatment discrimination in sport organizations. According to Ilgen and Youtz (1986), such discrimination can take place with regard to highly visible and tangible rewards, such as salary increases, or in a more subtle manner, such as exclusion from specific work groups. Indeed, previous research on the coaching profession has shown that persons holding minority status are often treated differently with regard to the allocation of rewards, the availability of resources, and opportunities for training and advancement (Cunningham & Sagas, 2004; Knoppers et al., 1989; Knoppers, Bedker-Meyer, Ewing, & Forrest, 1990, 1991; Lovett & Lowry, 1994; Stangl & Kane, 1991).

Given the presence of access and treatment discrimination within the sport context, it is necessary to understand the effects of such discrimination. Outcomes associated with access discrimination are somewhat straightforward, in that persons are denied access to positions within organizations or professions (Ilgen & Youtz, 1986). However, because treatment discrimination can be subtle within the organizational context and takes place in many forms, there is less literature associated with the outcomes of such differential treatment (Ilgen & Youtz, 1986). As such, the proposed model attempts to look beyond simply identifying the occurrence of treatment discrimination by focusing on one specific manifestation (i.e., performance evaluation bias) and its effect on performance. Further, even though treatment discrimination has been studied across multiple disciplines, it has rarely been examined from an integrative or multidisciplinary viewpoint. As such, a model is proposed that comprises the integration of literature from numerous realms of the social sciences (organizational psychology, social psychology, human resource management, and sociology). The applicability of this model will be demonstrated
within the sport context and focus on one particular form of treatment discrimination: differential performance appraisals.

The purpose of this article is to propose an integrative model that specifies how treatment discrimination, as manifested in performance evaluation bias, can influence employee performance. Specifically, it is expected that, as a result of the categorization process (i.e., the process by which people sort the self and others into in-groups and out-groups), persons who are different from the rater may receive lower performance appraisals. As a result, these low performance appraisals may lead to fewer organizational resources, rewards, and opportunities, thus potentially resulting in self-limiting behavior and ultimately decreased job performance. Potential moderators of these relationships are also discussed. An illustrative summary of the model is presented in Figure 1. The theoretical framework and specific propositions are presented in the following sections.

Theoretical Framework

Social Identity Theory

Social identity theory holds that people, in an attempt to simplify the complicated world around them, classify themselves and others into various social categories (Ashforth & Mael, 1989; Tajfel & Turner, 1979). Within group contexts, individuals base such classifications on salient characteristics such as age, ethnicity, gender, religion, organizational membership, and so on. To the extent that identification with a specific social category results in high levels of self-esteem, an individual may maximize group differences through positive evaluation of like individuals (i.e., in-group members) or negative evaluation of dissimilar individuals (i.e., out-group members) (Cunningham, 2004c; Roberson & Block, 2001). Indeed, these differential evaluations strengthen the “us” and “them” dynamic that results from the categorization process, subsequently heightening the potential for intergroup bias.

The literature has demonstrated that the categorization process often plays an important role in demographically diverse organizations. For example, Williams and O’Reilly (1998), after their review of 40 years of diversity research, noted that “it is clear that there are potentially negative consequences from social categorization processes operating in groups” (p. 119). The “negative consequences” to which Williams and O’Reilly refer are the propensity for persons to classify others who are demographically different as out-group members. Consider, for example, the case of a sport organization with a Caucasian manager. To the extent that ethnicity is a salient category within that context, the Caucasian manager is likely to consider ethnic minorities as out-group members. Thus, as seen in the first portion of Figure 1, demographic cues are likely to lead to the categorization of others into in-groups and out-groups.

The placement of persons into in-groups and out-groups based on their demographic similarity, or lack thereof, is important because of the benefits associated with in-group status. Within the organizational context, research suggests that members of in-groups and out-groups experience work differently (Cunningham, 2004c; Fink et al., 2001; Perdue, Dovidio, Gurtman, & Tyler, 1990). In-group members are allocated more rewards, given greater autonomy, included in participative
Figure 1 — Illustrative summary of conceptual model.
Discrimination and Performance Evaluations

539

decision making, granted greater access to organizational information, presented with challenging work tasks, and given higher levels of supervisory support than their out-group counterparts (Gerstner & Day, 1997; Greenhaus et al., 1990; Ilgen and Youtz, 1986; Perdue et al., 1990; Sagas & Cunningham, 2004). Furthermore, as a result of these benefits, in-group members experience higher levels of job performance, job satisfaction, career success, organizational commitment, and performance evaluations than do out-group members (Gerstner & Day, 1997; Greenhaus et al., 1990; Ilgen and Youtz, 1986; Sagas & Cunningham, 2004). Conversely, out-group members experience fewer positive work outcomes and higher instances of differential treatment within the organizational setting due to their status (Ashforth & Mael, 1989; Button, 2001; Ellemers, De Gilder, & Haslam, 2004; Gaertner & Dovidio, 2000). Of particular importance to this investigation is the differential treatment that emerges from performance evaluation bias. The importance of performance appraisals and evidence of bias in this process is discussed in the following section.

Performance Evaluations

MacLean and Chelladurai (1995) suggest that the success of organizations and their employees is dictated by the accuracy of feedback gained from effective performance evaluations. Judge and Ferris (1993) also note that supervisor ratings of subordinate performance are critically influential to an organization’s human resource functions, actions, and outcomes. These authors argue that “there is perhaps not a more important human resource system in organizations than performance evaluations” (p. 80). Indeed, performance evaluations serve as a means to select, train, retain, motivate, reward, and develop employees (Cunningham & Dixon, 2003).

Whereas performance evaluations have been identified as fundamental for the success of organizations and employees, they are also extremely susceptible to bias (Cunningham & Dixon, 2003; Ilgen & Youtz, 1986). At the organizational level, performance evaluations are suggested to be influenced by the level to which the rater trusts in the performance evaluation system as well as influenced by the level of organizational commitment held by the rater (Tziner, Murphy, Cleveland, Beaudin, & Marchland, 1998). On an individual level, similarities and differences between the rater and ratee can result in positive and negative performance evaluation bias, respectively. Indeed, this is consistent with social identity theory (Ashforth & Mael, 1989; Tajfel & Turner, 1979), in that persons who are different from the rater are likely afforded out-group membership and the accompanying negative consequences—in this case, poor evaluations—of such membership. Conversely, when an employee is similar to the rater, that employee is placed in the in-group and enjoys the benefits thereof, such as favorable performance evaluations. Empirical evidence supports this rationale, as rater–ratee similarity has been shown to be associated with high performance evaluations. These findings have been shown to hold across a variety of demographic characteristics, including ethnicity (Elvira & Town, 2001), sex (Tsui & O’Reilly, 1989), job tenure (Judge & Ferris, 1993; Tsui & O’Reilly, 1989), and age (Judge & Ferris, 1993). Further strengthening the presence of such bias and consistent with social identity theory is Buda, Reilly, and Smither’s (1991) work, which demonstrated that raters focused on information that was congruent with initial categorizations of ratees.
While the extant literature has demonstrated performance evaluation bias based on demographic differences, social identity theory holds that there are additional mechanisms to consider. Specifically, the current model extends this literature by explicitly articulating the presence of a mediating variable in the aforementioned relationship: the categorization and stereotyping process. Such articulation is more in line with social identity theory’s predictions and argues that the formation of in-groups and out-groups is the most proximal antecedent to differential performance evaluations. Therefore, as illustrated in Figure 1, the following proposition is made.

Proposition 1: As a result of demographic dissimilarity and the subsequent categorization process, out-group members are likely to receive lower performance evaluations than their in-group counterparts.

Self-Limiting Behavior

If demographic similarity, or the lack thereof, and the subsequent categorization process do lead to differential performance evaluations, it is next necessary to consider the results of such effects. Expanding on Ilgen and Youtz’s (1986) investigation of the evaluation and development of minority members within organizations, two outcomes of differential performance evaluations exist: lost opportunities and self-limiting behaviors. The lost opportunities effect is comprised of two main categories: role model presence and workgroup composition. Role model presence refers to the professional and psychological benefits accrued from organizational mentor–protégé relationships and sponsorship opportunities (Ilgen & Youtz, 1986). The workgroup composition component refers to the additional opportunities presented to in-group members (i.e., similar to the rater) of a particular workgroup. When compared to their out-group counterparts, in-group members may experience the allocation of more rewards, greater autonomy, inclusion in participative decision making, greater access to organizational information, and the presentation of challenging work tasks (Gerstner & Day, 1997; Greenhaus et al., 1990; Ilgen and Youtz, 1986; Perdue et al., 1990; Sagas & Cunningham, 2004).

The second outcome resulting from differential performance evaluations between in-group and out-group members is termed self-limiting behavior (Ilgen & Youtz, 1986). Ilgen and Youtz (1986) describe self-limiting behavior as long-term results of lost opportunities and further identify two main components: ability and motivation. Specifically, self-limiting behavior will emerge when feedback results in a lack of opportunities and compromises negative information with regard to one’s ability and motivation to perform a task. These processes are outlined in greater detail below.

Ability. Within the organizational context, the ability to perform at high levels is often dependent upon being offered challenging work tasks, advancement opportunities, and skill development opportunities. If employees are denied access to meaningful work and the chance to take part in training initiatives, their performance will suffer over time, especially in relation to persons who are not deprived of such work experiences. Thus, denial of these opportunities results in an individual’s inability to perform at high levels. To the degree that an organization uses results from performance evaluations to assign developmental opportunities (see
Discrimination and Performance Evaluations

Cunningham & Dixon, 2003; MacLean & Chelladurai, 1995), those receiving negative evaluations will not be granted the same opportunities as those receiving positive evaluations—thus highlighting the significance of evaluation bias. Therefore, when negative evaluations are given on the basis of something other than performance (e.g., demographic dissimilarity), such practices are likely to negatively affect the ability of the persons subjected to such bias.

Motivation. Exposure to negative feedback information, as well as denial of opportunities, may also result in a decrease in motivation toward a task. Simply put, an individual who is exposed to negative performance evaluation feedback may consciously or unconsciously put forth less effort. Further, this effect on motivation may be exacerbated when both exposure to negative feedback and denial of opportunities are repeated. Consciously, negative feedback may prompt one to no longer seek advancement opportunities owing to a perceived lack of qualifications and expectations of failure, thus limiting oneself (Ilgen & Youtz, 1986). Unconsciously, a person may internalize the negative feedback to the degree that he or she accepts such information and behaves in a confirmatory manner without realizing it, again, limiting him- or herself. Regardless of whether a decrease in motivation occurs consciously or unconsciously, performance detriment can occur, thus resulting in poor performance evaluations.

The concept of self-limiting behavior is highly related to self-efficacy, as both encompass an ability component as well as a motivational component. Whereas some level of self-limiting behavior may or may not emerge after performance evaluation information has been given, some level of one’s self-efficacy is present prior to performance as well as following the presentation of performance feedback information. Indeed, this reasoning is based on Bandura’s (1986) suggestion that self-efficacy emerges from many different sources, such as, past performance, psychological state, vicarious experiences, and verbal persuasion, any of which can be present prior to task performance. Judge and Bono’s (2001) meta-analysis of self-evaluation traits, job satisfaction, and job performance supports this notion by having identified self-efficacy as a fundamental determinant of job performance.

Self-efficacy is defined as one’s assessment of his or her capabilities toward successfully performing a task (i.e., ability component; Bandura, 1986). Further, an individual’s level of self-efficacy can impact one’s response to feedback information and subsequent performance (i.e., motivational component; Bandura, 1986). Research suggests that the degree to which an individual internalizes negative information affects his or her self-efficacy (Nease, Mudgett, & Quinones, 1999). Bandura (1986) suggested that repeated negative feedback tends to lead to lower self-efficacy. In a similar vein, Vance and Colella (1990) suggested that repeated exposure to negative feedback may result in a decreased effort and task withdrawal. Taken together, this line of research indicates a linkage between one’s self-efficacy and self-limiting behavior. As such, and in drawing from these investigations, the proposed model suggests that continued bias in performance evaluations can lead to a decrease in ability, motivation, and self-efficacy as well as performance decrements—that is, self-limiting behavior (Ilgen & Youtz, 1986). Specifically, the following proposition is made.

**Proposition 2:** Performance evaluation bias resulting in negative feedback toward out-group members will be positively related to self-limiting behavior.
Conversely, performance evaluation bias resulting in positive feedback toward in-group members will be negatively related to self-limiting behavior.

**Performance**

Job performance is the result of several factors, including ability, the motivation to perform the task, personality, and environmental factors (Barrick, Stewart, & Piotrowski, 2002; Colquitt, LePine, & Noe, 2002; Mount & Barrick, 1995). As such, the ability and motivational decrements associated with self-limited behavior (Ilgen & Youtz, 1986) are thought to have a significant influence on the actual performance of the employee. Consider for example, that if an employee has a decrease in ability resulting indirectly from the evaluation biases, that person will be less equipped to handle the day-to-day activities associated with the job. As a result, performance will likely suffer. Similarly, if one lacks the motivation and/or self-efficacy (i.e., perceived ability) to complete a task—as a result of the continued bias—performance again will decrease. Indeed, as argued above, Bandura's (1986) research in the area of self-efficacy, as well as Judge and Bono’s (2001) meta-analysis of the self-evaluation traits as they relate to job satisfaction and job performance, supports these tenets. Therefore, Proposition 3 is made as follows.

**Proposition 3:** Self-limiting behavior will be negatively related to work performance.

**Performance Feedback**

Poor performance can have several negative effects, such as decreased wages, termination, and the like. Here, two outcomes specifically related to the model are discussed: stereotype confirmation and poor performance appraisals. Each of these is discussed in detail below.

As previously noted, differential performance evaluations of persons demographically different from the rater are conceivably the result of the categorization process and the formation of in-groups and out-groups (Ashforth & Mael, 1989; Tajfel & Turner, 1979). The categorization process results in the stereotyping of in-group and out-group members (Tsui & Gutek, 1999), such that positive stereotypes are affixed to in-group members and less positive (or negative) stereotypes are attributed to out-group members. Therefore, in drawing from social identity theory to explain differential performance appraisals, biases in the evaluation process are potentially the result of stereotypes about the performance of out-group members. If, as the previous propositions hold, biases in performance evaluations ultimately lead to decreased performance, then the poor performance of out-group members will simply serve to reinforce the stereotypes of poor out-group performance formed by the rater. Indeed, this linkage leads to the development of Proposition 4.

**Proposition 4:** As a result of poor performance, initial stereotypes of out-group members will be confirmed, resulting in further categorizations of individuals of dissimilar demographic characteristics as out-group members.

In addition, another outcome of the poor performance resulting from biases and self-limiting behavior is continued low performance evaluations. Specifically, the following proposition is made.
Proposition 5: Poor performance will continue to result in poor performance evaluations of out-group members. Consequently, the possibility for further self-limiting behaviors also exists, thereby creating a process cyclical that is in nature.

Moderators

Thus far, direct effects have been specified in the first five propositions. In addition to these direct effects, it is also possible that moderators influence these relationships. The specification of moderators is useful in theory and model development because it specifies when and under what conditions the relationships are thought to take place (Bacharach, 1989). Four potential moderators are presented below next: task interdependence, time, social category identification, and a common group identity.

Task Interdependence

One possible moderator of the relationship between demographic dissimilarities and the categorization process is that of task interdependence. In the organizational setting, task interdependence is described as the extent to which a team or group member’s job, task, or project requires interaction, collaboration, and coordination with other group members with regard to activities and the exchange of information and materials (Doherty & Chelladurai, 1999; Timmerman, 2000; Van der Vegt & Van de Vliert, 2005). Indeed, it is through this increase in interpersonal contact and knowledge exchange that team members may begin to form personalized relationships with fellow team members, subsequently resulting in an appreciation of individual differences. Empirical evidence has suggested that task interdependence can influence team-related outcomes, such as team commitment and team satisfaction (Campion, Papper, & Medsker, 1996; Van der Vegt, Emans, & Van de Vliert, 2000). Campion et al. (1996) demonstrated that, through repeated and prolonged personal interaction, team members developed a sense of camaraderie and friendship, thus fostering feelings of satisfaction and commitment toward the team. However, within the context of diverse groups, theory and research suggests that task interdependence is related to poorer outcomes (Doherty & Chelladurai, 1999; Timmerman, 2000). For example, Doherty and Chelladurai (1999) argue that the effects of diversity, whether positive or negative, would be stronger in groups with high task interdependence. Presumably, high levels of task interdependence reinforce differences (because of the continuous interactions among diverse people), thereby strengthening whatever effects diversity may have on subsequent outcomes. This relationship has been supported empirically in the context of athletic teams. In an examination of effects of team composition (i.e., age and racial diversity) on team performance, Timmerman (2000) found that tasks that required high levels of interdependence (i.e., basketball) had a greater negative effect on the relationship between team composition and team performance than did tasks which required lower levels of interdependence (i.e., baseball).

In drawing from the aforementioned literature, it is proposed in the current model that when a rater and ratee work closely and/or are members of a common workgroup or team, task interdependence will moderate the relationship between demographic dissimilarities, the categorization process, and the reliance on
stereotypical information. As a result, instances of differential performance evaluations toward demographically dissimilar ratees and subsequent self-limiting behaviors may be strengthened. Specifically, the following hypothesis with regard to performance evaluations is made.

**Proposition 6a:** Tasks requiring high levels of interdependence between a rater and a ratee will increase categorization bias and decrease the reliance on stereotypical information.

**Time**

Another possible moderator of the relationship between demographic differences and the categorization process is the issue of time. Specifically, demographic or outward differences between individuals are quickly interpreted and used to make judgments (Tsui, Eagan, & O’Reilly, 1992). Harrison, Price, Gavin, and Florey (2002) refer to such difference as surface-level characteristics. As such, when dissimilar individuals spend very little time with one another, little can be learned beyond such visible differences as age, sex, ethnicity, cultural background, and so on. Consistent with social identity theory (Ashforth & Mael, 1989; Tajfel & Turner, 1979) and as stated previously, classifications and categorizations will emerge in relation to the self as a means of protecting one’s own social identity. As a result, an individual will positively evaluate individuals with similar surface-level characteristics (Ashforth & Mael, 1989; Harrison et al., 2002; Roberson & Block, 2001; Tsui & Gutek, 1999).

Just as there is a natural affinity toward individuals with similar surface-level characteristics, so too do individuals prefer to interact with those who share similar psychological or deep-level characteristics. Harrison et al. (2002) identify such characteristics as attitudes, values, beliefs, and personalities and suggest that similarities of such characteristics emerge through member interactions and over time. Empirical findings from these same authors indicated that whereas surface-level differences were important in the initial stages of the group, deep-level differences become more influential as the life of the group progressed. A similar pattern is expected with regard to interactions between raters and ratees, such that, through the passage of time, the effects of surface-level differences will be reduced. This rationale leads to the formation of the following.

**Proposition 6b:** The greater the amount of time allowed for rater-ratee interactions, the greater the likelihood that a rater will categorize beyond surface-level characteristics and begin to categorize based on deep-level characteristics. In turn, this will lessen the effect of demographic dissimilarity on the categorization of persons into in-groups and out-groups.

**Social Category Identification**

The degree to which an individual identifies with a particular subgroup is another potential moderator between demographic differences and the categorization process. This is particularly important as the salience of various demographic attributes differs for each individual and across different contexts (Cunningham, 2004c; Tsui & Gutek, 1999). Tsui and Gutek (1999) suggest that an individual’s
core identity emerges from a primary or highly important demographic category, further suggesting that, when appropriate, categorizations of others are based on the salience of the same category. However, if others are not highly distinct on that demographic category, alternative characteristics might be used. For instance, if a young African American male coach enters a sport organization of predominately older Caucasian male coaches, the relative importance of the coaching profession might be overlooked if demographic differences such as age and ethnicity are highly salient and used for social categorization. However, should the older Caucasian males hold the coaching profession to be the core of their identities, at least within that context, the age and ethnicity of the young African American coach should be of less importance when making such categorizations. Therefore, as illustrated in Figure 1, the degree to which an individual identifies with a particular social category is expected to influence the relationship between rater–ratee dissimilarities and the categorization process.

**Proposition 6c:** The degree to which a rater identifies with a specific social category (e.g., group membership) will influence the degree to which a rater categorizes a demographically dissimilar ratee as an out-group member on the basis of demographics alone.

**Common Group Identity**

A final moderator specified in the current model is the presence of a common group identity. As previously noted, in-group members receive numerous organizational benefits and experience more positive work outcomes than their out-group counterparts (Gerstner & Day, 1997; Greenhaus et al., 1990; Ilgen and Youtz, 1986; Perdue et al., 1990; Sagas & Cunningham, 2004). As suggested by Proposition 1, it is further expected that out-group members will receive lower performance evaluations as a result of their status. Consequently, a potential moderator between the categorization process and differential performance appraisals that occur between in-group and out-group members is the establishment of a common group identity through the process of recategorization.

Recategorization, the basis of Gaertner and Dovidio’s (2000) common in-group identity model, is a process by which in-group and out-group members in a given entity come to conceive of themselves as one single entity, thus eliminating intergroup bias. Specifically, rather than differentiating between in-group and out-group members, all members of a group are viewed as in-group members—the effects of which may be visible in the elimination of performance evaluation bias toward former out-group members. If recategorization takes place, the only rating differences that occur among members of the newly formed aggregate should be based upon actual performance differences rather than intergroup bias, as all persons would be considered in-group members and afforded in-group bias (see also Cunningham, 2004c). Therefore, the final expectation, as illustrated in Figure 1, is presented as follows.

**Proposition 7:** The presence of a single, common in-group identity will minimize and potentially eliminate negative performance evaluation bias toward former out-group members on the basis of out-group status, subsequently reducing the likelihood for self-limiting behavior and poor performance.
Discussion

Although previous research in sport management and sport sociology has documented the occurrence of both access and treatment discrimination, systematic examination of outcomes associated with these differential experiences has not been undertaken. The purpose of this article was to address this void. Specifically, in drawing from social identity theory (Ashforth & Mael, 1989; Tajfel & Turner, 1979), this model outlines how treatment discrimination, as manifested in performance evaluations, may ultimately lead to performance decrement among those facing such biases. Indeed, these biases are particularly relevant to the sport and leisure context, as most managers may be Caucasian, Protestant, able-bodied, heterosexual males (Fink et al., 2001), indicating that persons with characteristics different from these are likely to face the negative effects of performance evaluation bias. The final section discusses the implications, limitations, and future directions associated with this framework.

Implications

There are several potential implications of this proposed model, the first of which may shed some insight on previous empirical findings within the sport industry. As Ilgen and Youtz (1986) note, one result of self-limiting behavior is that persons may not seek advancement opportunities owing to a perceived lack of qualifications and expectations of failure. Extrapolating from this notion, the current model proposes that self-limiting behavior may result from demographic dissimilarity, performance evaluation bias, and negative feedback toward out-group members. Thus within the sport setting, it can be suggested that people who are demographically different from majority members of sport organizations may not seek advancement into upper-level positions as a result of self-limiting behavior. Recent research in the area of coaching supports this linkage, as Cunningham and Sagas (2002; Cunningham, Sagas, & Ashley, 2003) have found that female assistant coaches have less intent to become a head coach than do their male counterparts. In further support of the proposed model, Cunningham et al. found that gender differences in self-efficacy could at least partially explain these differences, as women had less coaching self-efficacy than did men. Because self-efficacy is an aspect of lost opportunities and to some extent self-limiting behavior, these findings are consistent with the current model and Ilgen and Youtz. Therefore, in drawing from the current model, one explanation for the gender differences in intentions to become a head coach could be the self-limiting behavior of women assistant coaches—behavior that results from biases and treatment discrimination.

A second implication of the current model refers to its practical application to the organizational setting. Specifically, the model specifies one moderator—the presence of a common identity—that could be used to reduce the effects of performance evaluation bias. Managers could use this information to implement strategies to reduce the negative effects of demographic dissimilarity. In building upon Gaertner, Mann, Murrell, and Dovidio’s (1989) suggestion that the formation of a common group identity results in more positive evaluations of former out-group members, further research has identified numerous recategorization strategies, each with the focus of creating a single common identity. Specifically, suggestions have
been made that, through the promotion of employee interaction and identification of a common fate, common threats, and common goals, managers can foster a common in-group identity, subsequently reducing evaluation bias (Allport, 1954; Tsui & Gutek, 1999; Rothgerber, 1997; Williams & O’Reilly, 1998). Therefore, when applied to the current model, the extent to which evaluation bias leads to lost opportunities, the self-limiting behavior and poor performance of out-group members will be mitigated by the formation of a common in-group identity.

Thirdly and also related to the applicability of the current model to the organizational setting, in addition to encouraging a common group identity, the current model calls attention to the potential need for managers to also take steps to remove some of the biases in performance evaluation systems. In turn, removing these biases can be seen through coworker satisfaction and workgroup preferences. Cunningham (2004b) suggested that performance appraisals of out-group members serve as a mediator between the formation of a common in-group identity and satisfaction with coworkers and preferences toward working with the group. Specifically, when the performances of former out-group members are perceived as similar to the performances of in-group members, coworker satisfaction and preferences to work with the group will be high. Extending upon this notion and applying it to the current model, it can further be suggested that enhanced rater perceptions of demographic dissimilar ratees as well as enhanced preferences to work with demographically dissimilar ratees will reduce the likelihood that performance evaluation bias will occur as a result of such dissimilarity.

Future Directions and Boundary Conditions

There are several avenues for future research stemming from this model. First, there is empirical evidence that demographic dissimilarities result in performance evaluation bias (Elvira & Town, 2001; Judge & Ferris, 1993; Tsui & O’Reilly, 1989). Poor performance evaluations, in turn, result in a number of work outcomes, including limited career success and opportunity (Greenhaus et al., 1990). It would be advantageous to examine whether limited career success and opportunity could also be attributed to the self-limiting behavior that the current model expects results in decreased performance. Additional research is needed to examine this possibility.

Perhaps one of the most fruitful avenues for future research stemming from this model is the testing of the proposed relationships. Given its vastness and complexity, the entire model may not be testable in a single study; however, portions of the model can most certainly be tested through experiments, surveys, and qualitative data collection. Previous research in related fields provides ideas for how this can be accomplished. For example, determining the effects of performance appraisals has been carried out through both experiments and large-scale surveys (for a comprehensive review, see Smither, London, & Reilly, 2005). Survey methods (e.g., Tsui & O’Reilly, 1989) and experiments (Cunningham, 2004b) have also been used to examine the linkage between demographic differences and subsequent work outcomes, thus suggesting that these methods may be best suited for testing many of the proposed relationships in the current model. For example, in testing the moderation effect of time on the relationship between demographic dissimilarities and categorizations, a method similar to that used in Harrison, Price, &
Bell’s (1998) investigation of the impact of time on the effects of surface-level and deep-level diversity on workgroup cohesion can be used. Specifically, Harrison et al. administered a series of questionnaires to two groups, over both a 3-week and 6-week time period, respectively, to assess the relative importance of surface-level and deep-level diversity among workgroup members. Results demonstrated that the more time members of a workgroup spent interacting with and sharing information about one another, the less important surface-level differences became. In turn, deep-level characteristics (values, attitudes, beliefs, etc.) became increasingly important among workgroup members. Indeed, similar methods can be applied to the current model to assess the influence of time on demographic dissimilarities and their relationship with categorizations and stereotypes within the rater–ratee dyad (see Proposition 6b).

As stated above and demonstrated through previous empirical investigations, both experiments and survey collection methods could be used in testing Propositions 2–5 as well as the moderating effects of Propositions 6 and 7. Interestingly, a review of the literature shows that qualitative methods are largely underutilized in performance evaluation research (Smither et al., 2005). Certainly, in-depth interviews or the use of focus groups could prove useful in qualitatively examining the propositions espoused in the current article. Future research should consider adopting any of these research techniques, or some combination thereof, to test the model’s propositions.

Additionally, the current model specifies a single outcome of the categorization process: bias in the performance evaluation process. However, there are additional outcomes as well, some of which have been examined and some of which have not. For example, social identity theory (Ashforth & Mael, 1989; Tajfel & Turner, 1979) and the relational demography perspective (Tsui & Gutek, 1999) put forward that persons different from their supervisors or work colleagues may experience work poorly. Thus, at the supervisor–subordinate dyad level, demographic dissimilarity might be related to outcomes such as low supervisor support, poor mentoring behaviors, and lack of career sponsorship. At the group level, being different from group members might be associated with poor group integration, low job satisfaction, and job stress. As all of these outcomes are associated with well-being (or lack thereof) in the workplace, and future research is needed to examine these relationships.

While the current model focuses on the process of categorization at both the individual (i.e., on the basis of demographic dissimilarities) and at the group level (i.e., formation of a common in-group identity), it should be noted that the relationship between categorization and performance evaluation bias may be moderated by numerous other factors not identified in the current model. Specifically, performance evaluation bias toward out-group members may be mitigated by the use of 360-degree feedback systems, rater training, and the establishment of specific performance measurement criteria and procedures. Indeed, research suggests such factors are influential when attempting to optimize performance appraisal systems (Atwater, Waldman, & Brett, 2002; Cunningham & Dixon, 2003; Fink & Longenecker, 1998; Lee, 1985). However, the use of such methods within the current model would require further investigation.

The current model may also be influenced by a number of temporary and permanent contextual or situational factors that may influence the categorization
process as well as the activity of rating subordinates. Indeed, Judge and Ferris (1993) identified the importance of examining the social and situational factors that shape rater cognition when seeking to fully understand the performance appraisal process. In line with this suggestion, performance ratings within the current model may be influenced by such contextual factors as rater opportunity to observe ratee performance (Judge & Ferris, 1993), rater knowledge of required job tasks (Lee, 1985), and perhaps even times of organizational and environmental uncertainty. Further, and of particular relevance to the sport organization context, categorizations and subsequent ratings may be influenced by job-specific stereotypical characteristics necessary to perform. For instance, as mentioned previously, the typical intercollegiate athletic employee is a White, Protestant, able-bodied, heterosexual male (Fink et al., 2001). As such, it could be suggested that these characteristics are congruent with perceptions of attributes that all individuals within the sporting profession should possess. Those individuals who deviate from this prototype may indeed be subject to differential performance ratings. However, the basis for such evaluation bias may be perceived demographic dissimilarity between the job type or position and the individual rather than the rater and the individual. Further investigation of this possibility is necessary both within and outside of the sport organization context.

Additionally, it must be acknowledged that although demographic dissimilarity may exist between a rater and a ratee, negative performance evaluations may be warranted if performance is truly poor. Indeed, this is the true function of the performance evaluation process. However, while research suggests that individuals who are performing poorly may initially respond to negative feedback with increased effort, this may not be the case when individuals are confronted with repeated exposure to negative feedback (Campion & Lord, 1992). As mentioned above, Bandura (1986) suggested that negative feedback may result in a decrease in self-efficacy. Expanding upon this, Nease et al. (1999) demonstrated that repeated exposure to negative feedback resulted in a continued detriment to self-efficacy. In building upon these findings and consistent with the current model, this decrease in self-efficacy (i.e., perceived ability) may result in self-limiting behavior, continued poor performance, and poor performance evaluations. Thus, the presence of self-limiting behavior may exist when accurate and warranted negative performance evaluation information (i.e., on the basis of poor performance), is provided.

**Conclusion**

The purpose of this article is to propose an integrative model that specifies how treatment discrimination may influence subsequent employee performance. Specifically, it is hypothesized that, as a result of the categorization process, persons who are different from the rater will receive lower performance appraisals and as a result, engage in self-limiting behavior and ultimately decreased performance. It is further hypothesized that this process could be influenced by the level of task interdependence (Campion et al., 1996; Doherty & Chelladurai, 1999; Dovidio, Kawakami, & Gaertner, 2000; Timmerman, 2000; Van der Vegt & Van de Vliert, 2005), the amount of time that raters and ratees spend interacting with one another (Harrison et al., 2002), the level of identification with a particular social category.
(Tsui & Gutek, 1999), and the recategorization process (Gaertner & Dovidio, 2000). This model has implications for the management of diverse organizations across numerous contexts. Additionally, future directions of inquiry and boundary conditions are offered, thus suggesting the next logical step is to empirically test the applicability of the model. Future research will address this possibility.

Acknowledgments

In appreciation of their constructive feedback and input, a special thank-you is extended to George B. Cunningham, Janet Fink, and the three anonymous reviewers of this manuscript.

References


