Journal of Sport & Exercise Psychology, 2007, 29, 608-628
© 2007 Human Kinetics, Inc.

Development and Validation of the Moral Disengagement in Sport Scale

Ian D. Boardley and Maria Kavussanu
University of Birmingham

A sport-specific measure of moral disengagement was developed in 2 studies. In Study 1, a 59-item questionnaire was developed and tested with 308 athletes from 5 team sports. A series of confirmatory factor analyses (CFA) testing different models suggested the model that best fitted the data had 6 first-order factors that could be represented by 1 second-order factor. Study 2 involved 305 athletes from the same 5 sports. CFA confirmed the 6-factor, second-order structure for the final 32-item measure. Results from Study 2 supported the construct validity of the scale, providing evidence for the factorial, concurrent, convergent, and discriminant validity. The Moral Disengagement in Sport Scale (MDSS) is proposed as a valid and reliable measure of moral disengagement for use in the sport context.

Key Words: sport, moral behavior, confirmatory factor analysis

Sport can bring joy and achievement to those who participate; at the same time, it suffers from the ills of those who break the rules and display aggression toward other participants. In a recent study of 803 youth-sport participants, nearly 10% confessed to cheating, 13% admitted trying to injure an opponent, 31% acknowledged arguing with an official, 13% acknowledged having made fun of a less-skilled teammate, and 27% disclosed acting like "bad sports" (Shields, Bredemeier, LaVoi, & Power, 2005). In addition to self-reported behaviors, research has documented the prevalence of observed negative social behaviors in sport. In their study of male soccer teams, Kavussanu and colleagues (Kavussanu, Seal, & Phillips, 2006) found that antisocial behaviors such as retaliating to a bad tackle were far more frequent and diverse than prosocial ones during soccer matches. An important goal of sport morality research is to understand what leads players to engage in both antisocial and prosocial action while playing sport.

The present study seeks to extend previous research on sport morality and is guided by the social cognitive theory of moral thought and action (Bandura, 1991). This theory considers various factors, including the consequences of the action, in defining behavior as moral or reprehensible. Bandura (1999) has also described morality as having dual aspects: inhibitive and proactive. The inhibitive aspect is the power to refrain from acting inhumanely. The proactive aspect is the power to behave humanely. In sport, the inhibitive aspect would be manifested when a player

The authors are with the School of Sport and Exercise Sciences, University of Birmingham, Edgbaston, Birmingham, United Kingdom.

refrains from injuring an opponent. The proactive aspect would be evidenced by a player helping an injured opponent. In past sport research (Kavussanu et al., 2006; Sage, Kavussanu, & Duda, 2006), the terms *prosocial* and *antisocial behavior* have been used to refer to the proactive and inhibitive aspects of morality, respectively. Prosocial behaviors have been defined as actions intended to benefit one or more persons other than oneself (Eisenberg & Fabes, 1998). In contrast, antisocial behaviors have been defined as behaviors intended to harm or disadvantage another (Kavussanu, 2006; Sage et al., 2006). High levels of morality are evidenced when one engages in prosocial behavior and refrains from engaging in antisocial actions (Bandura, 1999).

Team-sport athletes are subjected to numerous pressures to engage in transgressive acts. For example, when coaches and parents convey an expectation of success at all costs, players could be tempted to break the rules of the game in order to satisfy this expectation. Also, competing for key positions on a team can lead to athletes' feeling the need to transgress. Players are often evaluated based on the outcomes of their actions rather than the means through which they achieve them. For example, a rugby player in a defensive position with a record of very few missed tackles is likely to be commended for his or her performances even if these were achieved through questionable means. Finally, the intimidation or incapacitation of key opponents can also be a motivator for malevolent conduct. Though players may at times submit to these pressures and engage in reprehensible conduct, they are likely to do so through means that allow them to preserve their self-regard.

Bandura (1991) described a process through which moral conduct is regulated. When an individual engages in moral behavior, he or she may feel guilt or pride depending on the nature of the behavior. These affective reactions result from a process of self-monitoring and judgment regarding the moral nature of our actions. For example, a player may feel guilty having committed a foul on an opponent and having seen the opponent's injuries, or feel proud having helped an injured opponent. These feelings regulate future behavior anticipatorily, by reducing motivation for actions that result in negative affect, and increasing motivation for behaviors that result in positive feelings.

Although this process is thought to regulate moral conduct, individuals do not always act the way they should. Bandura (1991) suggests that people are able to violate personal standards without self-sanction through the selective use of eight psychosocial maneuvers, collectively known as *mechanisms of moral disengagement.* Moral disengagement reduces the inhibitive aspect of morality by acting as a moderator between transgressive behavior and the affective self-reactions that regulate behavior. This results in reducing or eliminating these reactions, thereby decreasing subsequent constraint on future transgressive behavior. Moral disengagement mechanisms operate on four aspects of harmful conduct and have been grouped into four sets that correspond to these aspects (Bandura, 1991).

The first set acts on the harmful behavior itself, reclassifying it as respectable and includes moral justification, euphemistic labeling, and advantageous comparison. *Moral justification* involves the cognitive reconstrual of blameworthy behaviors into honorable ones, making transgressive behavior personally and socially acceptable by portraying it as a facilitator of a valued social or moral purpose (Bandura, 1999). In sport, foul play and cheating could be justified as a way of protecting teammates, team honor, or reputation. *Euphemistic labeling* involves the selective

use of language that cognitively disguises culpable activities as less harmful (Bandura, 1999). Examples of euphemistic labeling in sport are when players talk of "bending the rules" when they really break them or "letting off steam" when they in fact act aggressively. *Advantageous comparison* involves comparing transgressive behaviors with more reprehensible activities, making them appear benevolent or trivial (Bandura, 1999). An example in sport is comparing the use of aggressive language to physical violence. Such methods allow players to convince themselves that their actions are inconsequential, thereby avoiding self-condemnation.

The second set of mechanisms targets the responsibility for action, aiming to reduce the accountability felt by an individual, and includes displacement and diffusion of responsibility. *Displacement of responsibility* occurs when people view their actions as resulting from social pressures or the directives of others and not something for which they are personally responsible (Bandura, 1999). By displacing responsibility to an authority figure such as the coach, athletes can employ unfair tactics that they would not usually contemplate. *Diffusion of responsibility* is achieved through division of labor, group decision making, or group action (Bandura, 1991). In division of labor, members of a group perform subdivided tasks that are considered harmless individually but when combined are harmful. Group decision making takes advantage of the fact that when everyone is responsible, no one actually feels personally responsible (Bandura, 1999), and in group action moral control is reduced through attributing any harm done to others within the group (Zimbardo, 1995). An example of diffusion of responsibility in sport is when all members of a team are involved in decision-making processes regarding the use of antisocial practices, thereby sharing the liability and reducing the personal accountability felt by each individual.

The third set targets the consequences that result from detrimental conduct and involves only one mechanism. *Distortion of consequences* occurs when an individual either avoids or cognitively minimizes the harm caused by his or her reprehensible action, thus weakening potential self-deterring reactions (Bandura, 1999). Previous research on obedient aggression demonstrated that people were less likely to continue detrimental conduct when the suffering of the victim became more apparent (Milgram, 1974); at that stage it may have been more difficult for the individual to avoid or minimize the harm caused. Use of this mechanism in sport is seen when players avoid finding out the extent of injuries sustained by opponents or deny the seriousness of the injuries of which they are aware.

The final set operates on the victim of the act and contains the dehumanization and attribution of blame mechanisms. *Dehumanization* involves cognitively depriving opponents of human qualities or attributing animalistic qualities to them (Bandura, 1999). The similarity one feels with another has an effect on how he or she treats that person. Cognitively accentuating differences between oneself and another allows one to treat that person badly without self-censure. When athletes describe opponents as animals or suggest they lack feelings they may in fact be dehumanizing them. *Attribution of blame* occurs when people see themselves as blameless victims driven to detrimental conduct by forcible provocation (Bandura, 1999). Through this process, transgressive behavior becomes excusable. In sport this occurs when players retaliate and blame their victim for their behavior.

Researchers have considered the impact of moral disengagement on moral behavior in various contexts and have developed instruments to measure this

construct in these contexts (e.g., Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Osofsky, Bandura, & Zimbardo, 2005). The first empirical study to investigate this issue found that moral disengagement was positively linked to delinquent behavior and proneness to aggression and negatively related with prosocial behavior (Bandura et al., 1996). Other studies have established links between moral disengagement and transgressive acts in society (Bandura, Caprara, Barbaranelli, Pastorelli, & Regalia, 2001), implementation of the death penalty (Osofsky et al., 2005), transgressing civic duties (Caprara & Capanna, 2005), support of military action (McAlister, 2001), support of military force (McAlister, Bandura, & Owen, 2006), and bullying in schools (Menesini et al., 2003) and prisons (South & Wood, 2006). To date, only one study has investigated moral disengagement in a physical activity context. Specifically, Lucidi and colleagues (Lucidi, Grano, Leone, Lombardo, & Pesce, 2004) evaluated whether moral disengagement was related to intention to use doping substances in Italian high school students. The researchers reported a significant positive association between moral disengagement and intention to use doping products.

Moral disengagement in the studies described above has been measured using various instruments. In these instruments, different factor structures have been identified. Scales measuring moral disengagement toward interpersonal aggression (Bandura et al., 1996), civic duties (Caprara & Capanna, 2005), and support of military action (McAlister, 2001), have displayed a unidimensional structure. In contrast, scales examining moral disengagement by those involved in capital punishment (Osofsky et al., 2005) and in support of military force (McAlister et al., 2006) have displayed a four-dimensional structure. The four dimensions in the scale measuring moral disengagement in relation to capital punishment corresponded to the four aspects of detrimental conduct targeted by the four sets of moral disengagement mechanisms; the four-factor solution represents the core of the conceptual model described by Bandura (1991). Thus, to date research has reported either a single or a four-factor structure for moral disengagement measures. However, based on Bandura's (1991) theorizing, one should also consider the construct of moral disengagement to consist of eight interrelated dimensions representing the eight mechanisms. Thus, the possibility of the presence of eight factors should be considered when developing a measure of moral disengagement.

The links between moral disengagement and detrimental conduct in different contexts highlight the need for developing a measure of moral disengagement specific to sport. There are several arguments that reason for the development of such a measure. First, moral disengagement mechanisms refer to specific behaviors. Thus, sport behaviors must be included in the items to allow examination of levels of disengagement in the sport context. Second, the established links between moral disengagement and transgressive behavior in society suggest that investigation of moral disengagement in sport may aid our understanding of moral conduct in sport. Finally, previous research into moral disengagement has involved the development of context-specific instruments, suggesting that moral disengagement may be context specific (e.g., Bandura et al., 1996; Caprara & Capanna, 2005; Osofsky et al., 2005). In sum, effective measurement of moral disengagement in sport requires the development of a sport-specific instrument.

The purpose of the studies presented here was to develop a sport-specific measure of moral disengagement, termed the *Moral Disengagement in Sport Scale*

(MDSS). Based on Bandura's (1991) theorizing, we hypothesized that the scale would have an eight-factor structure. Also, because of the related nature of the mechanisms, we expected that the eight factors would be subsumed under a higher-order factor. However, based on past research (Bandura et al., 1996; Osofsky et al., 2005) we also tested the possibility that the scale had either a one or four-factor structure. Study 1 included the preliminary scale development and confirmation of the factor structure. Study 2 confirmed the factor structure identified in Study 1 with an independent sample and provided evidence for the construct validity of the MDSS.

Study 1

The aim of Study 1 was to develop a scale that would reflect the multidimensional nature of moral disengagement. The methods employed in this study were consistent with those proposed by Haynes, Richard, and Kubany (1995) and Clark and Watson (1995). Initially, a list of items representing the eight mechanisms was developed. Next, sports psychology experts were asked to comment on the face and content validity of these items. Then, amendments were made based on these comments, and the scale was pilot-tested. Finally, CFA was used to test the fit of the factor structure against the one-, four-, and eight-factor models using a large heterogeneous sample of team-sport athletes. Details of these procedures are reported in the following sections.

Method

Preliminary Scale Development

Initially, we developed a large pool of 59 items intended to measure the eight moral disengagement mechanisms (Bandura, 1991). A small number of items (n = 12) were developed by adapting items from the scale of Bandura et al. (1996) to a team-sport context. For example, the dehumanization item "Some people deserve to be treated like animals" became "Some opponents deserve to be treated like animals." Most items (n = 47) were developed after consultations with sport psychologists, team coaches, and active sportspersons.

Next, the content validity of the items was examined. Content validity is concerned with whether instrument items are characteristic of the domain they are intended to measure (Kline, 2005) and is an important aspect of scale development (Haynes et al., 1995). The most effective way of examining content validity is through expert opinion. To determine content validity, the complete set of items was evaluated by 10 sport psychology professionals experienced in scale development. The experts were provided with the definitions of the eight moral disengagement mechanisms (Bandura et al., 1996) and with the list of items, and were asked to (a) rate how representative each item was of the respective definition on a 5-point Likert scale ranging from −2 (*not at all representative*) to +2 (*very*

representative), and (b) comment on each item's relevance to sport. Based on the median values recorded and the qualitative assessments of the items, adjustments were made to the wording of 31 items, 2 items were removed, and 2 new items were developed. The content validity of the new items was confirmed by six sport psychology professionals.

Finally, the measure was pilot-tested with a sample of sport and exercise science students who were active team-sport participants ($N = 105$) in order to (a) ensure that items were correlated ($r > .15$) with other items within each mechanism and with the total scale score, (b) test the difficulty of the items, and (c) obtain feedback about their wording (see Clark & Watson, 1995). Following this procedure, minor adjustments were made to a small number of items. A 7-point Likert scale anchored by 1 (*strongly disagree*) and 7 (*strongly agree*) was used in the pilot testing as well as in all subsequent data collections because this format offers the best compromise between reliability, validity, discriminating power, and respondent preferences (Preston & Colman, 2000).

Participants

Participants were representative of both genders ($n_{male} = 191$; $n_{female} = 117$), and their age ranged from 12 to 68 years ($M = 21.73$, $SD = 8.18$). They were drawn from teams playing at the competitive club and university levels in the regions of central and northern England. Respondents participated competitively in the sports of soccer ($n = 108$), netball ($n = 61$), hockey ($n = 49$), rugby ($n = 64$), and basketball ($n = 26$). They had played their main team sport competitively for an average of 9.70 years ($SD = 7.07$) and had played for their current team for an average of 3.60 years ($SD = 4.21$). The large variety of sports, ages, and number of years involved in sport was targeted specifically to ensure heterogeneity within the sample (see Clark & Watson, 1995).

Procedure

Following clearance by the ethics committee of a British university, we contacted the head coaches of 32 teams from the relevant sports regarding participation of their athletes in the sutdy. All coaches agreed to allow their athletes to participate. Data from all participants were collected during designated training sessions that were scheduled a minimum of 1 day before a game or match. One of the investigators visited all teams and distributed questionnaires to the athletes. Before completing the questionnaire, all respondents were informed that the survey examined sporting attitudes, that honesty in responses was vital to the success of the study, and that they should complete the questionnaire with their main competitive team sport in mind. It was also explained that all responses would be kept strictly confidential and would be used only for research purposes. Participants signed an informed consent form prior to completing the questionnaire, which took approximately 10–15 min to complete. In total, 326 questionnaires were distributed, with a response rate of 94%, resulting in a sample of 308 athletes.

Results

Preliminary Analyses

Before testing the factor structure of the instrument, a two-stage process was used to select the items most effective at measuring their respective mechanism and for use in subsequent model testing. First, interitem correlations were examined within each of the eight moral disengagement mechanisms; any item that was correlated less than .15 with the other items in its category was removed (Clark & Watson, 1995). Second, exploratory factor analysis (EFA) was conducted on each of the eight moral disengagement subscales using principal axis extraction. Each of these EFAs produced a unidimensional factor structure; extraction was based on the criterion of an eigenvalue of greater than 1. Subscales were examined individually in order to retain the best indicators of their underlying latent variable (Jöreskog, 1993). Based on these analyses, we selected for each mechanism the five items with the highest loadings on their respective factor. A total of 40 items with minimum factor loadings of .40 were selected and used in subsequent analyses.

Confirmatory Factor Analyses

The next step involved using CFA to test the fit of a series of models. As discussed previously, Bandura (1991) described eight mechanisms of moral disengagement, implying that the construct is multidimensional. Thus, we expected the scale to have eight dimensions. However, because past research has reported either a unidimensional or a four-factor structure, we also examined these possibilities. The four-factor structure involved the four sets of mechanisms that operate on the four aspects of detrimental conduct (Bandura, 1991). In the four- and eight-factor models, factors were specified to be correlated with each other because they are theoretically related. Confirmatory factor analysis was employed because (a) it offers a rigorous test of the plausibility of the factor structure, and (b) is the most appropriate method for confirming hypothesized factor structures (Fabrigar, Wegener, MacCallum, & Strahan, 1999).

The CFA analyses were conducted using the EQS 6.1 (Bentler & Wu, 2002) statistical package and its Maximum Likelihood method. In the first analysis, the normalized estimate of Mardia's coefficient of multivariate kurtosis was high (28.5), indicating substantial deviation from multivariate normality. Therefore, the Robust Maximum Likelihood estimation method was used in all subsequent analyses. This method produces more accurate standard errors, chi-squared values, and fit indices when the data are not normally distributed (Bentler & Wu, 2002). The case numbers with the largest contribution to normalized multivariate kurtosis suggested minimal impact of outliers. Therefore, no cases were deleted.

The indices used to test the fit of each model were the Satorra–Bentler scaled robust chi-square ($R\chi^2$), the robust comparative fit index (RCFI), the robust non-normed fit index (RNNFI), the standardized root mean square residual (SRMR), and the root mean square error of approximation (RMSEA). Because the $R\chi^2$ statistic is sensitive to sample size, the ratio of chi-square to degrees of freedom was also used to judge model fit. In "perfect" models, this ratio is 1.0, although ratios below 2.0 are considered acceptable. Values from RCFI and RNNFI greater than .90 indicate

acceptable model fit, whereas very good model fit is attained when the RCFI and the RNNFI values are close to .95, the RMSEA is less than .06, and the SRMR is less than .08 (Hu & Bentler, 1999). To compare models, the robust consistent Akaike information criterion (RCAIC) was used. When models are compared, the model with the lowest value is preferred (Hair, Anderson, Tatham, & Black, 1998).

The first analysis examined the fit of the eight-factor model theorized by Bandura (1991). Based on previous analyses (i.e., interitem correlations and EFA), five items were retained for each of the eight mechanisms and were specified as indicators of their corresponding factor in the model (M1a). The results showed an inadequate fit of the model to the data (Table 1, row 1) suggesting the need for respecification. Subsequently, 10 items that had large modification indices as indicated by the Lagrange multiplier (LM) test and/or large standardized residuals were removed in a series of CFAs. A final model (M1b) with 30 items produced an eight-factor solution with satisfactory fit indices (Table 1, row 2). However, the solution for this model included a parameter estimate that made it inadmissible. This was a Heywood case: The correlation between the moral justification and the euphemistic labeling factors was above 1 ($r = 1.01$), indicating either model misspecification or empirical underidentification (Kline, 2005). As the data were not empirically underidentified, this value suggested the possibility that moral justification and euphemistic labeling are not distinct factors. Following the merging of these factors, the resulting seven-factor model had an almost identical fit to the eight-factor model, thus supporting the merging of these two factors. The results of this analysis can be obtained from the first author.

Table 1 Summary of Fit Indices for All CFA Models

Model	*df*	Rχ^2	Rχ^2/*df*	RCFI	RNNFI	SRMR	RMSEA	RCAIC
			Study 1					
1. M1a, 40 items	712	1277.05	1.79	.87	.85	.06	.05	−3514.8
2. M1b, 30 items	377	621.81	1.65	.92	.91	.06	.05	−1915.4
3. M2a, 30 items	390	654.43	1.68	.91	.90	.06	.05	−1970.3
4. 2nd order, M2a	400	747.07	1.87	.89	.88	.06	.05	−1945.0
5. M3, 30 items	399	785.24	1.97	.87	.86	.07	.06	−1900.1
6. M4, 30 items	405	1042.13	2.57	.79	.77	.07	.07	−1683.6
			Study 2					
7. M2a, 30 items	390	597.72	1.53	.94	.94	.05	.04	−2023.2
8. M2b, 32 items	449	674.59	1.50	.95	.94	.05	.04	−2342.8
9. 2nd order, M2b	459	740.16	1.61	.93	.93	.05	.05	−2344.5

Note. df = degrees of freedom; Rχ^2 = Satorra–Bentler scaled chi-square; RCFI = robust comparative fit index; RNNFI = robust non-normed fit index; SRMR = standardized root mean square residual; RMSEA = root mean square error of approximation; RCAIC = robust consistent Akaike information criterion. M1 = 8-factor model; M2 = 6-factor model; M3 = 4-factor model; M4 = 1-factor model.

To examine whether any other factors were not empirically distinct, we tested a series of models that merged other *theoretically related* factors. The only other factors that could be merged *without* a reduction in fit were the diffusion and displacement of responsibility factors. The merging of these two factors to form a six-factor model resulted in a model fit that achieved acceptable values (Table 1, row 3), which were not significantly different from those for the seven-factor model. Comparisons between these models were based on the similarity of the RCAIC values and work that has shown that a value of ΔCFI smaller than or equal to −0.01 indicates measurement invariance between models (Cheung & Rensvold, 2002). If two models have similar explanatory power for the same data, the more parsimonious model is preferred (Kline, 2005). Thus, we accepted the six-factor model. No other modifications were made.

As described above, the six-factor model included two new factors. The first factor was termed *conduct reconstrual* and was formed by merging the moral justification and euphemistic labeling factors. The second factor was named *non-responsibility* and was created by combining the diffusion of responsibility and displacement of responsibility factors. The merging of the factors in these pairs has theoretical support: Moral justification and euphemistic labeling both act by cognitively reconstruing the conduct as less harmful, and diffusion of responsibility and displacement of responsibility both act by minimizing personal responsibility (Bandura, 1991). The six-factor model also included the advantageous comparison, distortion of consequences, dehumanization, and attribution of blame factors.

The six factors represent different dimensions of the *same* overriding construct: moral disengagement. Therefore, the factors were hypothesized to be related to each other. Indeed, medium-to-high correlations were observed among them (see Table 2). When a construct consists of related factors, it is important to examine whether the factor correlations can be explained by one or more higher-order factors (Kline, 2005). To investigate this possibility, hierarchical versions of M2a based on theory (Bandura, 1991) were also tested. We investigated whether the six first-order factors (a) could be represented by four factors (i.e., two second-order and two first-order factors) that corresponded to the four aspects of detrimental conduct, and (b) could be explained by only one second-order factor.

Table 2 CFA Factor Correlations of Six-Factor First-Order Model

Factor	1	2	3	4	5	6
1. Conduct reconstrual	—	.83	.77	.80	.91	.87
2. Advantageous comparison	.85	—	.70	.89	.73	.71
3. Nonresponsibility	.71	.65	—	.66	.72	.84
4. Distorting consequences	.68	.90	.51	—	.67	.73
5. Dehumanization	.87	.68	.56	.54	—	.80
6. Attribution of blame	.82	.65	.86	.54	.70	—

Note. Correlations below the diagonal are from Study 1; correlations above the diagonal are from Study 2. For all correlations, $p < .05$.

The first model resulted in nonsignificant parameter estimates and was rejected. In contrast, the model with *one* second-order factor had no problems with parameter estimates. Although the fit of this model (Table 1, row 4) was slightly worse than the fit of the corresponding first-order model (Table 1, row 3) and marginally failed to achieve an acceptable level, it was viewed as adequate to support a second-order structure. The fit of a second-order model cannot be better than the fit of the equivalent first-order one (Marsh, 1987). However, if the fit of the higher-order model approaches that of the first-order model, the hierarchical model should be preferred because it is more parsimonious. This second-order factor was named *sport moral disengagement*. Support for the second-order model was shown by an RCAIC value very similar to that of the first-order model. The RCAIC was deemed more suitable than the chi-square difference test for comparing the fit of these two models because the latter test is sensitive to sample size (Cheung & Rensvold, 2002).

The next model tested was the four-factor model reported by Osofsky et al. (2005). The four factors in this model represent the four sets of mechanisms that target the different aspects of detrimental conduct. The 30 items used in previous models were entered into a four-factor model (M3) and were specified to load on their corresponding factors. This model (Table 1, row 5) had a less-than-adequate fit. Further modifications of this model using all 40 items failed to achieve an acceptable fit, confirming that this model is not representative of the actual factor structure of this scale. The final model tested was the unidimensional model (see Bandura et al., 1996). The items were entered into the model so that all 30 were specified to load onto a single factor. This model (M4) indicated a very poor fit (Table 1, row 6). Further modifications of this model utilizing all 40 items failed to achieve an acceptable fit, suggesting that this model is also not representative of the actual factor structure of this scale.

In summary, the 30-item six-factor model produced the best fit to the data, outperforming both the four-factor and one-factor models. The second-order, six-factor model produced results similar to those of the first-order model and achieved an acceptable level of fit for all but the RCFI and RNNFI indices. Thus, this model was selected as the best model because of its greater parsimony (Kline, 2005). Although the removal of some items achieved an acceptable fit for the first-order model, the possibility that the performance of certain items was sample specific (Fabrigar et al., 1999) led to 38 items from the original 59 being retained in the second study. These were the 38 best-performing items from Study 1.

Gender, Age, Sport Type, and Moral Disengagement

In this study we used male and female athletes from five sports, representing a wide age range. We examined whether athletes differed on moral disengagement as a function of gender, sport type, and age. ANOVA indicated gender, $F(1, 306) = 85.74$, $p < .001$, and sport type, $F(4, 303) = 22.78$, $p < .001$, differences on sport moral disengagement. Men displayed higher levels of moral disengagement than women ($M = 3.66$ vs. 2.91), and rugby ($M = 3.70$) and soccer ($M = 3.70$) players had higher levels of moral disengagement than basketball ($M = 3.14$), hockey ($M = 3.10$), and netball ($M = 2.80$) players. Finally, age was negatively correlated with moral disengagement ($r = -.32$, $p < .001$).

Study 2

The first study provided evidence that the hierarchical model with six first-order factors and one second-order factor offered the most satisfactory fit for the MDSS. However, because the results of CFA can be sample specific, it is recommended that data from an independent sample are tested with CFA to confirm factor structures (Fabrigar et al., 1999). The first purpose of the second study was to confirm the factor structure identified in the first study. A second purpose was to examine the criterion-related validity of the instrument through the provision of evidence relating to concurrent, convergent, and discriminant validity. Details of the procedures used to examine these purposes are outlined below.

Method

Participants

Participants were representative of both genders (n_{male} = 217; n_{female} = 88), ranged in age from 12 to 55 years (M = 21.81, SD = 9.53), and were drawn from teams playing at the competitive club and university levels in the regions of central and northern England. Respondents participated competitively in the sports of soccer (n = 125), netball (n = 75), hockey (n = 12), rugby (n = 49), and basketball (n = 44). They had played their main team sport competitively for an average of 9.43 years (SD = 7.14) and participated in their current team for an average of 5.72 years (SD = 5.76).

Measures

Moral Disengagement in Sport. Moral disengagement in sport was measured using the MDSS developed in Study 1, utilizing 40 items: the 38 of the original 59 items that performed best in Study 1, plus the two new items developed to ensure that the instrument had an equal number of items representing each mechanism (i.e., 5 per mechanism). The new items were "Fighting is okay if it is done to protect a teammate," measuring moral justification, and "Teasing an opponent does not really hurt him/her," measuring distortion of consequences. We chose to develop two new items rather than use items developed for Study 1 because of the poor performance of items intended to measure these two mechanisms in the first study. Sport psychology professionals (N = 6) were again consulted regarding the new items and confirmed that they had acceptable content and face validity.

Even though a total of 40 items were used in data collection to measure moral disengagement in sport, similar to Bandura et al. (1996), we intended to produce a scale in which each mechanism would be measured using only 4 items, thereby resulting in a 32-item scale. Thus, we included the extra eight items with the intention to utilize them if the two new items or the items contained in the final six-factor model, identified as the best model in Study 1, performed poorly in the CFAs conducted in Study 2.

Moral Disengagement in Society. The 32-item measure of moral disengagement developed by Bandura et al. (1996) was used to measure participants' degree of societal moral disengagement. This instrument was included to test the convergent validity of the MDSS. Evidence for convergent validity is established if a scale is shown to correlate at least moderately with established measures of the same or similar constructs (Kline, 2005). Example items from this scale are "A kid in a gang should not be blamed for the trouble the gang causes" (diffusion of responsibility) and "If people are careless where they leave their things it is their own fault if they are stolen" (attribution of blame). Items were rated on a 7-point scale anchored by 1 (*strongly disagree*) and 7 (*strongly agree*). This scale has demonstrated satisfactory internal consistency with a Cronbach's (1951) alpha coefficient of .82 (Bandura et al., 1996). The alpha coefficient in the present study was .94.

Prosocial and Antisocial Behavior in Sport. A 14-item measure of reported prosocial and antisocial behavior in sport adapted from past research (Kavussanu, 2006) was used to indicate the frequency of such behaviors. Prosocial and antisocial behaviors were measured to test the concurrent validity of the MDSS. Concurrent validity is concerned with whether a measure can predict a theoretically related external criterion when scores on the predictor and criterion are collected at the same time (Kline, 2005). The scale was originally developed for use in soccer; therefore, some items were removed (e.g., diving to fool the referee) or altered (i.e., tried to get an opponent booked was changed to tried to get an opponent penalized by an official). Owing to the relatively low level of internal consistency ($\alpha = .68$) for the prosocial behavior subscale of the original instrument, three new items were developed (i.e., returned the ball to the opposition, apologized to an opponent after fouling him/her, encouraged an opponent). Face and content validity of the new items was confirmed by sport psychology experts. Participants were asked to report how often they had engaged in each of the 14 behaviors during the season on a 7-point scale anchored by 1 (*never*) and 7 (*very often*).

Past research has reported a two-factor structure for this scale using principle components analysis (Kavussanu, 2006). In this study, CFA was used to confirm this factor structure. The scale achieved very good levels of fit, $R\chi^2$ (76) = 118.67, $R\chi^2/df = 1.56$, CFI = 0.96, NNFI = 0.95, RMSEA = 0.04, SRMR = 0.06. The correlation between the two factors was −.18 ($p < .05$). The two-factor structure reflected antisocial behavior (8 items, $\alpha = .86$) and prosocial behavior (6 items, $\alpha = .74$).

Procedure

The head coaches of 51 teams participating in the relevant sports were contacted regarding their athletes' participation in the study. All coaches agreed to allow their athletes to participate. Participant contact and data collection procedures and instructions were identical to those used in Study 1. In total, 318 questionnaires were distributed, with a response rate of 96%, resulting in a sample of 305 athletes.

Results

Confirmation of Factor Structure

In line with the a priori approach of Study 2, CFA was performed on three predetermined models: (a) the 30-item, six-factor, first-order model confirmed in Study 1; (b) the same model with the two new items added; and (c) a hierarchical model with six first-order factors and one second-order factor. A high value for the normalized estimate of Mardia's coefficient of multivariate kurtosis (35.6) suggested that the Robust ML estimation method should again be used. The case numbers with the largest contribution to normalized multivariate kurtosis suggested minimal impact of outliers; therefore, no cases were removed. The results of these analyses are presented in Table 1.

The first model tested was the six-factor first-order, 30-item model (M2a). The results showed a good model fit (Table 1, row 7). The second model tested was the M2a with the two new items added (M2b), and achieved a better fit than the first model (Table 1, row 8). The factor correlations for the six first-order factors are presented in Table 2. Finally, the hierarchical version of M2b with one second-order factor was tested. As expected (see Marsh, 1987), the fit of the hierarchical model (Table 1, row 9) was worse than the fit of the first-order model (Table 1, row 8). However, the model still displayed a good fit and RCAIC values similar to those of the equivalent first-order model. Based on this, the hierarchical model was accepted. Factor loadings, error variances, and the final 32 items are presented in Table 3. Owing to the good fit indices attained by the 32-item model, there was no need to utilize in data analysis the extra eight items used in data collection.

Descriptive Statistics

Descriptive statistics and correlations among the subscales of the MDSS can be found in Table 4. In general, these indicate low-to-moderate levels of moral disengagement and moderate-to-high correlations among the subscales. This table also presents the estimates of internal consistency for the MDSS subscales using Cronbach's (1951) alpha. These coefficients ranged from .73 to .95, indicating acceptable to very good levels of reliability.

Construct Validity

The second purpose of Study 2 was to provide further evidence for the construct validity of the MDSS by examining its convergent, concurrent, and discriminant validity. To examine convergent validity, the correlation between societal (Bandura et al., 1996) and sport moral disengagement as measured by the MDSS was computed. Evidence for the convergent validity of the MDSS would be established if levels of societal moral disengagement correlated at a moderately high level with levels of sport moral disengagement. However, if the correlation is too high ($r > .90$), this would suggest that the new instrument may be redundant (Kline, 2005). Societal moral disengagement was positively related to sport moral disengagement, $r = .71$, $p < .01$.

Table 3 Items, Standardized Factor Loadings, and Error Variances (EV) for Final 32-Item Second-Order Model

Item	Loading	EV	Factor
1. It is okay to be hostile to an opponent who has insulted your teammate/s.	.65	.76	CR
2. It is okay for players to lie to officials if it helps their team.	.60	.80	CR
3. Fouling an opponent is okay if it discourages him/her from injuring your teammates.	.74	.67	CR
4. Fighting is okay if it is done to protect a teammate.	.68	.73	CR
5. Injuring an opponent is a way of teaching him/her a lesson.	.69	.73	CR
6. Bending the rules is a way of evening things up.	.68	.73	CR
7. Acting aggressively is just a way of showing you are a tough opponent.	.69	.73	CR
8. Arguing with officials is a way of keeping them on their toes.	.67	.74	CR
9. Mocking an opponent is not bad compared to injuring him/her.	.51	.86	AC
10. Compared to physical violence, verbally provoking an opponent is not that bad.	.77	.63	AC
11. Shouting at an opponent is okay as long as it does not end in violent conduct.	.82	.58	AC
12. Winding an opponent up is nothing compared to screaming abuse in his/her face.	.75	.66	AC
13. A player should not be blamed for using illegal tactics if everyone on the team is doing it.	.55	.84	NR
14. It is unfair to blame players who only play a small part in unsportsmanlike tactics used by their team.	.60	.80	NR
15. A team decision to use unsportsmanlike tactics is just that, and not the responsibility of any individual on the team.	.54	.84	NR
16. Players should not take responsibility for negative consequences of their actions if they are following team decisions.	.64	.77	NR
17. A player is not responsible for acting aggressively if this is encouraged by his/her parents.	.60	.80	NR
18. A player should not be blamed for injuring an opponent if the coach reinforces such behavior.	.60	.80	NR
19. If players are not disciplined for unsportsmanlike conduct they should not be blamed for behaving this way.	.58	.82	NR
20. A player should not be blamed for arguing with officials if he/she has seen the coach doing it.	.61	.80	NR
21. Mocking an opponent does not really hurt him/her.	.77	.64	DC
22. Insults among players do not really hurt anyone.	.74	.67	DC
23. Aggressive language toward an opponent does not actually harm anyone.	.81	.58	DC
24. Teasing an opponent does not really hurt him/her.	.80	.60	DC

(continued)

Table 3 *(continued)*

Item	Loading	EV	Factor
25. Some opponents deserve to be treated like animals.	.71	.70	DH
26. It is okay to treat badly an opponent who behaves like an animal.	.85	.53	DH
27. If an opponent acts like an animal he/she deserves to be treated like one.	.73	.68	DH
28. If an opponent does not act humanly he/she should be made to suffer.	.61	.79	DH
29. If a player is mocked by an opponent, it is the opponent's fault if the player then tries to injure him/her.	.61	.80	AB
30. Players who get mistreated have usually done something to deserve it.	.62	.79	AB
31. If a player retaliates to something an opponent has done, it is the opponent's fault.	.64	.77	AB
32. A player should not be held responsible if he/she retaliates when fouled.	.71	.71	AB
First-order factors			
1. Conduct Reconstrual	.95	.33	SMD
2. Advantageous Comparison	.90	.44	SMD
3. Nonresponsibility	.82	.58	SMD
4. Distortion of Consequences	.85	.52	SMD
5. Dehumanization	.88	.48	SMD
6. Attribution of Blame	.89	.45	SMD

Note. Mocking an opposition player means *verbally* taunting him or her to cause distraction or provoke a punishable reaction. Winding up an opposition player means *physically or verbally* taunting him or her to cause distraction or provoke a punishable reaction. CR = conduct reconstrual; AC = advantageous comparison; NR = nonresponsibility; DC = distortion of consequences; DH = dehumanization; AB = attribution of blame; SMD = sport moral disengagement.

Table 4 Study 2: Correlations, Reliability Estimates, and Descriptive Statistics for the MDSS Constructs (*N* = 305)

Scale	1	2	3	4	5	6	7
1. Conduct reconstrual	(.86)						
2. Advantageous comparison	.70	(.81)					
3. Nonresponsibility	.64	.56	(.81)				
4. Distortion of consequences	.69	.74	.54	(.86)			
5. Dehumanization	.78	.60	.59	.57	(.81)		
6. Attribution of blame	.69	.56	.63	.57	.64	(.73)	
7. Sport moral disengagement	.90	.84	.77	.83	.84	.81	(.95)
M	3.20	3.75	2.95	3.13	3.31	3.46	3.22
SD	1.25	1.45	1.07	1.42	1.42	1.32	1.05
Range	1.00–6.00	1.00–6.67	1.00–6.38	1.00–7.00	1.00–6.55	1.00–6.75	1.00–5.87

Note. Scores can range from 1 to 7. Sport moral disengagement values are average of all 32 items. For all correlations, $p < .01$. Alpha coefficients are reported on the diagonal.

Concurrent validity was assessed by measuring the associations between sport moral disengagement and antisocial and prosocial behavior in sport. Past research (Bandura et al., 1996) suggests that moral disengagement should correlate positively with antisocial behaviors and negatively with prosocial behaviors. Concurrent validity of the MDSS would be established if such relationships were found in this study. Sport moral disengagement was positively related to antisocial behavior ($r = .60$, $p < .01$) and inversely linked to prosocial behavior ($r = -.34$, $p < .01$). Weaker relationships were observed between societal moral disengagement and antisocial ($r = .42$, $p < .01$) and prosocial ($r = -.31$, $p < .01$) behavior.

Discriminant validity of the MDSS subscales was examined by evaluating (a) the correlations among the factors representing the six subscales and (b) the correlations between the subscales and prosocial and antisocial behavior. The factor intercorrelations ranged from .51 to .90 in Study 1 and from .66 to .91 in Study 2, suggesting that some subscales are clearly distinct from others, whereas some share a considerable amount of variance with others (see Table 2). For example, in both studies, advantageous comparison was distinct from dehumanization, nonresponsibility, and attribution of blame; nonresponsibility was distinct from conduct reconstrual, distortion of consequences, and dehumanization; and the last two were distinct from each other and from attribution of blame. In contrast, conduct reconstrual was highly related to advantageous comparison, dehumanization, and attribution of blame; advantageous comparison was highly related to distortion of consequences; and attribution of blame was highly related to nonresponsibility. The overall pattern of factor correlations indicated that some, but not all, subscales were clearly distinct from others, thus demonstrating reasonably good levels of discriminant validity for some subscales but low levels of discriminant validity for others.

Examination of the correlations of the subscales with prosocial and antisocial behavior suggests that *some* subscales have unique predictive capabilities. For

example, the correlation between nonresponsibility and antisocial behavior (r = .40, $p < .01$) was clearly lower than that between conduct reconstrual and antisocial behavior (r = .66, $p < .01$). Conduct reconstrual also correlated more highly with prosocial behavior ($r = -.35$, $p < .01$) than did advantageous comparison ($r = -.25$, $p < .01$) and nonresponsibility ($r = -.26$, $p < .01$). However, other subscales had very similar correlations with the outcome variables. For instance, advantageous comparison and distortion of consequences were similarly related to antisocial (r = .54 and r = .56, respectively) and prosocial ($r = -.25$ and $r = -.29$, respectively) behavior. Thus, our results provided evidence that some—but not all—subscales have unique predictive capabilities providing support for discriminant validity for some of the MDSS subscales.

Gender, Age, Sport Type, and Moral Disengagement

In this study we again used athletes representing both genders from five sports, representing a wide age range. Similar to the first study, ANOVA indicated gender, $F(1, 303) = 61.46$, $p < .001$, and sport type, $F(4, 300) = 21.66$, $p < .001$, differences on sport moral disengagement. Males displayed higher levels of moral disengagement than females (M = 3.55 vs. 2.53), and rugby (M = 3.72) and soccer (M = 3.69) players had higher levels of moral disengagement than basketball ($M = 2.84$), and netball (M = 2.50) players. Finally, age was again negatively related to moral disengagement ($r = -.22$, $p = < .001$).

Discussion

The use of moral disengagement has been documented in several contexts—society, prisons, and schools. Moral disengagement plays an important role in antisocial behavior in these contexts (Bandura et al., 1996, 2001; Menesini et al., 2003; Osofsky et al., 2005; South & Wood, 2006). Although antisocial behaviors also occur in sport (e.g., Kavussanu et al., 2006; Shields et al., 2005), to date no study has investigated moral disengagement in the context of sport. The present research sought to fill this gap in the literature by developing a sport-specific measure of moral disengagement.

Bandura (1991) has described eight mechanisms that people use to morally disengage. Accordingly, we developed items and tested a model with eight dimensions. However, results from both studies revealed that moral disengagement in sport, as measured by the MDSS, is best conceptualized as having six dimensions. Specifically, two pairs of factors were merged based on the CFA results: Moral justification and euphemistic labeling formed the conduct reconstrual factor, and displacement and diffusion of responsibility formed the nonresponsibility factor. The merging of these two pairs of factors is consistent with theory (Bandura, 1991): The mechanisms that formed each pair disengage moral restraint from detrimental conduct by acting on the *same* aspect of the conduct. This suggests that they have a *similar* function. This similarity may be large enough to make the mechanisms in each of the two pairs empirically inseparable. In support of this, Bandura and colleagues (Osofsky et al., 2005) have also reported a factor that included diffusion and displacement of responsibility as well as a factor that included moral justification and euphemistic labeling.

Following the confirmation of the six first-order factors, the hierarchical nature of the MDSS was examined. Results supported the presence of a second-order sport moral disengagement factor, indicating that all mechanisms are part of one overriding construct. This finding is consistent with Bandura's (1991) theorizing that the moral disengagement maneuvers are different ways of achieving the *same* goal: to disengage moral restraint from detrimental conduct and minimize self-reproach for such actions. The presence of a higher-order factor increases the utility of the MDSS, allowing it to be used as either a measure of general moral disengagement in sport, or as a measure of the six subscale constructs.

This is the first research to report a six-factor first-order structure of moral disengagement. Other moral disengagement scales have displayed either a single (e.g., Bandura et al., 1996) or a four-factor (e.g., Osofsky et al., 2005) structure. We can offer two explanations for the inconsistency in the number of factors reported between the present work and past research. First, it is likely that the factor structure of moral disengagement is context specific; that is, the way in which moral disengagement operates depends on the specific environment in which it occurs. Second, past research has examined the factor structure of measures using EFA and has not always reported the full details of these analyses (e.g., Bandura et al., 1996). In the present studies we used CFA to test the factor structure of our scale. Conducting CFA on other scales may indicate the presence of a structure similar to that found in the present work. Despite the different factor structures identified in different measures, most studies have successfully used the total scale score in their analyses, suggesting that the presence of one overriding construct is common in most contexts.

The construct validity of the MDSS was demonstrated by providing evidence for concurrent and convergent validity for the overall scale, and support for the discriminant validity of *some* subscales. Concurrent validity was evidenced by the strong positive correlations between sport moral disengagement and antisocial behavior and the moderate negative correlations between sport moral disengagement and prosocial behavior. The ability of the MDSS to concurrently predict these theoretically related constructs supports its usefulness in future research. The convergent validity of the MDSS was evidenced by a strong positive association between sport and societal moral disengagement. Finally, discriminant validity was evident for some subscales—but not others—as indicated by the magnitude of the factor intercorrelations and the relationships of the subscales with prosocial and antisocial behavior.

The high correlations between some of the factors indicate substantial redundancy between the subscales assessing the mechanisms represented by these factors. The highest redundancy was observed between conduct reconstrual and both dehumanization and attribution of blame as well as between advantageous comparison and distortion of consequences. Clearly, these subscales share a considerable amount of variance with each other, indicating low levels of discriminant validity. Owing to the high amount of common variance, these subscales have limited ability to differentially predict outcome variables. Despite the high intercorrelations, we kept these subscales separate in the present studies because the mechanisms they measure operate at different aspects of detrimental conduct, as described by Bandura (1999). However, the issue of redundancy in these subscales is important and should be addressed in future research, as high correlations between subscales are problematic.

Sport moral disengagement was positively linked with antisocial behavior and negatively related to prosocial behavior. Specifically, sport moral disengagement was associated positively with behaviors such as trying to injure opponents and breaking the rules of the game and negatively with behaviors such as helping injured opponents and congratulating opponents for good play. These results are consistent with Bandura's (1991) views that moral disengagement increases the frequency of transgressive behaviors and decreases the frequency of benevolent ones and suggest that moral disengagement in sport may have implications for the inhibitive and proactive aspects of morality (Bandura, 1999).

Our findings provide other evidence for the suitability of our scale as a measure of moral disengagement in sport. First, the relationships between the MDSS and antisocial and prosocial behaviors were stronger than those between societal moral disengagement and these behaviors. This suggests that the context-specific measure developed in the present studies has greater predictive ability in the context of sport than has the original measure of moral disengagement. Second, the internal consistency of the scale was established. Finally, the high reliability of the single second-order dimension ($\alpha = .95$) supports the premise that the scale consists of subscales measuring constructs that are part of a higher-order construct.

Study 2 demonstrated that certain MDSS subscales were more strongly related to antisocial behaviors than others. Specifically, conduct reconstrual—composed of moral justification and euphemistic labeling—had the strongest association with antisocial behavior, and nonresponsibility—composed of diffusion and displacement of responsibility—had the weakest relationship. Previous research (Bandura et al., 1996) has also reported that the mechanisms targeting the harmful behavior (i.e., moral justification, euphemistic labeling, and advantageous comparison) or the victim (i.e., dehumanization and attribution of blame) were more strongly associated with aggressive behavior than were the diffusion and displacement of responsibility or distortion of consequences mechanisms. Thus, the relationships between certain subscales and behaviors may be consistent across contexts.

Although on average participants reported low-to-moderate levels of moral disengagement, a considerable proportion of athletes do disengage morally, as indicated by the range of responses. Recent qualitative research has corroborated the use of moral disengagement in sport. For example, Long and colleagues asked young elite athletes their reasons for transgression (Long, Pantaléon, Bruant, & d'Arripe-Longueville, 2006). Participants articulated reasons where mechanisms of moral disengagement were clearly evident. For instance, a male rugby player explained that if an opponent hurts a teammate and the referee does not punish the perpetrator he will "kill the opponent in the next play" (p. 342). Two moral disengagement mechanisms are evident here: morally justifying the behavior for the social purpose of protecting a teammate and attributing blame to the victim for initially fouling his teammate. Taken together with past research, our findings indicate that moral disengagement takes place in sport.

In sum, we developed a context-specific instrument, the MDSS, to measure moral disengagement in sport, and established evidence for the instrument's construct validity and reliability. Because of its multidimensional hierarchical structure, this instrument can be used to measure moral disengagement at the level of the six first-order subscales or at the level of the higher-order construct of sport moral disengagement. Our findings indicate that athletes use moral disengagement maneuvers when playing sport, and that this use is associated positively with

antisocial conduct and negatively with prosocial acts. These results suggest that investigating moral disengagement in sport has potential for enhancing our understanding of the social moral conduct that takes place in this context.

Future Research

The results from the two studies presented here provide evidence for the psychometric properties of the MDSS. However, validation is a continuing process (Clark & Watson, 1995) and future studies should continue to examine and improve the validity of the MDSS. In particular, future research should address the issue of low discriminant validity found for some of the MDSS subscales. One way to enhance low discriminant validity is to develop new items for each of the highly related subscales that are more distinct from the items in other subscales, yet reflect the content of the mechanism they are intended to measure. However, if the high correlations persist, researchers should consider merging the highly correlated factors and test the fit of models with fewer factors than the six identified in the present work. Research could also provide evidence for test–retest reliability, and examine the invariance of the factor structure across gender, sport type, and cultures using multisample analyses. The advent of the MDSS opens up the opportunity for further research that investigates this construct. Of particular interest would be research considering whether the various mechanisms can consistently predict moral behavior. We believe that the MDSS can be used in future research to enhance our understanding of the role of moral disengagement on moral behavior in the context of sport.

References

Bandura, A. (1991). Social cognitive theory of moral thought and action. In W.M. Kurtines & J.L. Gewirtz (Eds.), *Handbook of moral behavior and development: Theory, research, and applications* (Vol. 1, pp. 71-129). Hillsdale, NJ: Erlbaum.

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review, 3*, 193–209.

Bandura, A., Barbaranelli, C., Caprara, G.V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology, 71*, 364-374.

Bandura, A., Caprara, G.V., Barbaranelli, C., Pastorelli, C., & Regalia, C. (2001). Sociocognitive self-regulatory mechanisms governing transgressive behavior. *Journal of Personality and Social Psychology, 80*, 125-135.

Bentler, P.M., & Wu, E.J.C. (2002). *EQS 6 for Windows user's guide.* Multivariate Software.

Caprara, G.V., & Capanna, C. (2005). Moral civic disengagement and values. *Ricerche di Psicologia, 28*, 67-84.

Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness of fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233-255.

Clark, L.A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*, 309-319.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.

Eisenberg, N., & Fabes, R.A. (1998). Prosocial development. In N. Eisenberg (Ed.), *Handbook of child psychology. Vol. 3: Social, emotional, and personality development* (pp. 701-778). New York: Wiley.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. (1998). *Multivariate data analysis* (5th Ed.). Upper Saddle River, NJ: Prentice Hall.

Haynes, S.N., Richard, D.C.S., & Kubany, E.S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*, 238-247.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Jöreskog, K.G. (1993). Testing structural equation models. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 294-316). Newbury Park, CA: Sage.

Kavussanu, M. (2006). Motivational predictors of prosocial and antisocial behaviour in football. *Journal of Sports Sciences, 24,* 575-588.

Kavussanu, M., Seal, A., & Phillips, D. (2006). Observed prosocial and antisocial behaviors in male soccer teams: Age differences across adolescence and the role of motivational variables. *Journal of Applied Sport Psychology, 18*, 326-344.

Kline, R. (2005). *Principles and practice of structural equation modeling* (2nd Ed.). New York: The Guilford Press.

Long, T., Pantaléon, N., Bruant, G., & d'Arripe-Longueville, F. (2006). A qualitative study of moral reasoning of young elite athletes. *The Sport Psychologist*, *20*, 330-347.

Lucidi, F., Grano, C., Leone, L., Lombardo, C., & Pesce, C. (2004). Determinants of the intention to use doping substances: An empirical contribution in a sample of Italian adolescents. *International Journal of Sport Psychology, 35*, 133-148.

Marsh, H.W. (1987). The hierarchical structure of self-concept and the application of hierarchical confirmatory factor analysis. *Journal of Educational Measurement, 24*, 17-39.

McAlister, A.L. (2001). Moral disengagement: Measurement and modification. *Journal of Peace Research, 38*, 87-99.

McAlister, A.L., Bandura, A., & Owen, S.V. (2006). Mechanisms of moral disengagement in support of military force: The impact of Sept. 11. *Journal of Social and Clinical Psychology,* 25, 141-165.

Menesini, E., Sanchez, V., Fonzi, A., Ortega, R., Constabile, A., & Lo Feudo, G. (2003). Moral emotions and bullying: A cross-national comparison of differences between bullies, victims and outsiders. *Aggressive Behavior, 29*, 515-530.

Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper & Row.

Osofsky, M.J., Bandura, A., & Zimbardo, P.G. (2005). The role of moral disengagement in the execution process. *Law and Human Behavior, 29*, 371-393.

Preston, C.C., & Colman, A.M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica, 104*, 1-15.

Sage, L., Kavussanu, M., & Duda, J.L. (2006). Moral identity and motivational predictors of prosocial and antisocial functioning in association football players. *Journal of Sports Sciences, 24*, 455-466.

Shields, D.L., Bredemeier, B.L., LaVoi, N.M., & Power, F.C. (2005). The sport behavior of youth, parents and coaches. The good, the bad and the ugly. *Journal of Research in Character Education, 3*, 43-59.

South, C.R., & Wood, J. (2006). Bullying in prisons: The importance of perceived social status, prisonization, and moral disengagement. *Aggressive Behavior, 32*, 490-501.

Zimbardo, P.G. (1995). The psychology of evil: A situationist perspective on recruiting good people to engage in anti-social acts. *Research in Social Psychology, 11*, 125-133.

Manuscript submitted: December 22, 2006 *Revision accepted:* May 25, 2007