

Psychometric properties of the Multidimensional Sport Leadership Scale comparison to Multifactorial Leadership Questionnaire

A. RUI GOMES*, CLARA SIMÃES**, CATARINA MORAIS***,
and RUI RESENDE****

(*) *University of Minho, School of Psychology, Portugal*

(**) *University of Minho, School of Nursing, Portugal*

(***) *Research Centre for Human Development, Faculty of Education and Psychology, Universidade Católica Portuguesa, Porto*

(****) *University Institute of Maia, Portugal*

Clarifying the actions of leaders and reducing construct redundancy in evaluation leadership is vital to improve knowledge about leadership. This study compares the psychometric properties of Multidimensional Sport Leadership Scale (MSLS; Gomes & Resende, 2014) and Multifactor Leadership Questionnaire (MLQ-5X; Bass & Avolio, 2000). Data collection included 379 athletes competing at national level of collective sports. Athletes fulfilled the MSL, MLQ-5X, and the Athlete Satisfaction Questionnaire (ASQ; Riener & Chelladurai, 1998). Data analysis included construct validity, scale reliability, convergent validity, and discriminant validity. Construct validity revealed original structures of MSLS and MLQ-5X, although MLQ-5X presented higher problems of internal consistency (Cronbach's alpha) and of convergent validity than MSLS. Transformational constructs of both instruments revealed convergent validity with satisfaction with leadership of ASQ instrument but MLQ-5X transactional constructs assumed stronger correlation with ASQ than the MSLS transactional constructs. In conclusion, MSLS revealed better psychometric properties than MLQ-5X, representing an alternative to evaluate leadership.

KEY WORDS: Decision-making, Leadership, Transformational leadership, Transactional leadership, Sports coaches.

This study was conducted at the Psychology Research Centre (PSI/01662), School of Psychology, University of Minho, and supported by the Portuguese Foundation for Science and Technology and the Portuguese Ministry of Science, Technology and Higher Education (UID/PSI/01662/2019), through the national funds (PIDDAC).

Correspondence to: Rui Gomes, Universidade do Minho, Escola de Psicologia, Campus de Gualtar, 4710-057 Braga, Portugal (e-mail: rgomes@psi.uminho.pt).

The topic of leadership, particularly the one of transformational and transactional leadership made major influence on social and human sciences (Judge & Piccolo, 2004). In fact, understanding how a specific person produces influences on individuals, groups, organizations, and even societies is very challenging for professionals dedicated to study this phenomenon and interested in promoting positive leadership. Transformational-transactional leadership paradigm, more specifically the full range leadership theory from Bass (1985), represents one of the most used proposals on understanding how leaders can produce extraordinary results on others (Bass, 2008). The model divides leadership in two major areas (transformational and transactional) including also into the passive approach of *Laissez-Faire*, which is basically understood as the absence of leadership. The transformational leadership is the one related with better results achieved by leaders. It starts with the articulation and representation of a positive and challenging vision to team members in order to increase their efforts until a level where outstanding results can occur (Bass, 1995). In simple words, transformational leaders inspire followers to believe in a collective positive mission; this believe increases followers' commitment and contribution to the mission that, in turn, augments the chances of performance beyond expectations (Bass, 1985). Transactional leadership is based in an exchange relation between the leader that sets the tasks and goals to achieve, and the followers that agree to do the tasks in exchange of some material or psychological compensation (e.g., payment, awards, rewards, among others) (Bass, 1995). In order to evaluate these types of leadership, Bass and Avolio (1991, 2000) proposed the Multifactor Leadership Questionnaire (MLQ-5X). The instrument evaluates the full range leadership theory by including five transformational leadership factors, three transactional leadership factors, and one nonleadership or *Laissez-Faire* leadership factor (see also, Antonakis, Avolio, & Sivasubramaniam, 2003; Avolio & Bass, 2004). Transformational leadership includes Inspirational Motivation (tendency of the leader to articulate and represent a positive vision in order to provide meaning and challenge to followers' work), Idealized Influence-Attributed (tendency of the leaders be perceived as charismatic due their ability to develop trust and confidence of followers toward the mission), Idealized Influence-Behavior (tendency of the leader to stimulate and acting according the collective mission and values), Intellectual Stimulation (tendency of the leader to challenging the followers' beliefs and assumptions in order to think about new solutions to solve existing problems); and Individualized Consideration (tendency of the leader to consider the followers' individual needs and tendency to develop their individual strengths). Transactional leadership includes Contingent Reward (tendency

of the leader to provide rewards in exchange of followers' efforts related to accomplish the defined tasks); Active Management-by-Exception (tendency of the leader to actively monitor and acting if deviations from rules and standards may occur); and Passive Management-by-Exception (tendency of the leader to only intervene after detecting errors or if standards have not been met). Laissez-faire is the absence of the leader regarding their role and responsibilities by not acting when that was necessary (Bass & Avolio, 2000).

The impact produced by the full range leadership theory and MLQ-5X in literature is outstanding. In fact, the theory is one of most cited in theoretical proposals of leadership and the instrument is one of the most used in research (Hoch, Bommer, Dulebohn, & Wu, 2018). However, along the years some main concerns exists about the understanding and measurement of constructs included in the model and derived versions of MLQ (Antonakis, Bastardo, Jacquart, & Shamir, 2016). As recently referred by Banks, Gooty, Ross, Williams, and Harrington (2018) there is construct redundancy in defining and bounding leader behaviors, which makes very difficult to determine the impact produced by distinct theories of leadership on individuals and organizations. For example, the reviews provided by Banks et al. (2016) and Hoch et al. (2018) indicated construct redundancy between transformational leadership of the full range leadership theory and more recent approaches of leadership, as is the case of authentic or ethical leadership. These concerns are also applied to the MLQ instrument being reported a significant overlap among idealized influence (attributes and behavior) and inspirational motivation that addresses problems of construct definition and measurement of charismatic leadership (Antonakis et al., 2016; Carless, 1998). These indications raise the possibility of ameliorate the instrument, and by consequence the full range leadership theory. This study aligns this logic by introducing three topics of reflection and improvement in the measurement of leadership. The first and second topics are related to transformational and transactional leadership and the third topic is related with decision-making leadership.

Starting by transformational leadership, this domain depends mainly of the leader' ability to set a challenging and positive mission that should reflect a long-term vision of common-shared values (Rowold, 2006). Then, leaders have to inspire followers to give their best efforts in order to fulfill the mission (Bass, 1985). This transformational inspiration of leaders is very important because we may have an interesting mission without any substantial change on attitudes and behaviors of followers directed to concretize the mission; thus, assuring this relation in the evaluation of transformational leadership seems vital to capture the intrinsic nature of changes produced by

leaders. This relation between the vision and subsequent inspiration do not seem to be very clear in the MLQ-5X instrument. In fact, Idealized Influence includes the leaders' confidence, power, and influence over the followers but also the stimulation of a positive vision for followers; this same stimulation of optimism about the future is included in Inspirational Motivation when leaders articulate a compelling vision of the future or when they express confidence that goals will be achieved. Another aspect to consider is that during this inspirational process, leaders may have to confront technical and personal issues; that is, followers may not possess all the necessary knowledge and abilities to perform their tasks and the existence of positive and strong relations between the leader and followers may be crucial in order to augment the chances of followers' commitment toward the mission. It is worth of note that MLQ-5X Intellectual Stimulation and Individualized Consideration dimensions seems to appeal to technical and personal aspects of leader and followers' relationship. In fact, Intellectual Stimulation is somewhat related to tasks when leaders encourages followers to adopt original solutions to current problems and Individualized Consideration may implicate that leaders understand personal needs and expectations of follower in order to best adjust tasks to each one. However, neither Intellectual Stimulation or Individualized Consideration explain how leaders stimulates technical executions of tasks or how they promote positive and strong personal relations with followers. This is somewhat surprising because there is a long tradition of conceptualizing the leaders behaviors on task orientated (when the leader provides structures and pathways for followers to effectively perform the tasks) and people - or relationship - oriented leadership (when the leader focuses on stimulating and maintaining good relationships with followers by being considerate) (Sosik & Jung, 2018). Transformational leadership may not be equivalent with task and relationship leadership, representing a new form of leadership, but is quite difficult to conceptualize and understand the leader-follower relationships without any transformation produced by the leader on technical abilities of followers and without having strong and positive relations of confidence and mutual trust between both of them. In other words, producing changes on technical abilities of followers and building strong and confident relationships with followers may have enough transformational power to change attitudes, beliefs, and behaviors of followers that can, in consequence, augment their involvement toward the established mission. Some literature in transformational leadership reinforces this idea by revising the way these dimensions are evaluated. For example, Rafferty and Griffin (2004) hypothesized and tested a different factor structure for transformational leadership and conclude by the inclusion of five factors (vision,

inspirational communication, intellectual stimulation, supportive leadership, and personal recognition), one of them is very close to relationship orientated leadership (supportive leadership). Besides, DeRue, Nahrang, Wellman, and Humphrey (2011) in an integrative meta-analytic analysis concluded that trait and behavioral theories of leadership (where are included task and relationship orientated leadership) are still important for explaining variance in outcomes of leadership, such as leader effectiveness, follower job satisfaction, and group/unit performance (e.g., 20 to 70%).

Transactional leadership of the full range leadership theory includes Contingent Reward and both forms of management-by-exception. This means that leaders seem to assume actions when things happens as desired (in this case, the leader uses Contingent Reward) or when things deviate (or may deviate) from the leaders' expectations (in this case, leader uses Management-by-Exception). However, leaders may also respond negatively when followers do not complete the tasks or when they execute tasks in an improperly way and, as consequence, levels of performance are below the expectations. In this case, Contingent Punishment (or Negative Feedback) may be a possibility, existing indications that may not represent an irrelevant or non-significant behavior assumed by leaders (Aryee, Chen, Sun, & Debrah, 2007; Pelletier, 2010; Turnnidge & Côté, 2019). It is quite interesting that the MLQ-5X does not include this type of leadership, although Avolio and Bass (1995) assumed that punishment represents a transactional behavior because leaders can assume corrective, threatening, or disciplining actions when followers fail to meet a specific standard of performance. Accordingly, there is no reason to believe that leaders can respond to the performance and behaviors of followers in a negative way in an effort to prevent performance problems or erratic behaviors in the future. Thus, introducing, at least, the Negative Feedback in the pool of transactional behaviors may increase realism and specificity to what can happen in transactional exchanges between the leader and the followers. Besides this potential lack of dimensions included in transactional leadership, other indications seem to suggest the need of reformulating this domain. Specifically, there are indications of overlap between Passive Management-by-Exception and Laissez-faire (Hinkin & Schriesheim, 2008) and the possibility of Contingent Reward may be a distinct domain of leadership regarding Management-by-Exception (Goodwin, Wofford, & Whittington, 2001, Tejeda, Scandura, & Pillai, 2001).

Finally, the full range leadership theory and MLQ-5X does not provide any indications about how leaders make decisions. Decision-making refers to the way leaders manage their power reflected by the way they assume decisions related to important aspects of their activity. As the previous case of

task and relationship oriented leadership, for decision-making there is also a long tradition in the study of leadership to analyze how leaders assume actions. For example, Situational Leadership Theory proposes that leadership behaviors should vary according levels of competence and commitment assumed by followers, meaning that the leader' authority may diminishes as competence and commitment of followers augment (Blanchard, 2007). Besides, there is substantial indications that distinct types of decision-making produces differential effects on followers and teams (Lorinkova, Pearsall, Sims Jr, 2013; Oc, 2018), which make very difficult to ignore decision-making as an important factor contributing to the efficacy of leadership. It is quite interesting that Bass (1998) recognized the importance of decision-making on leadership but decided not to include it on both the model and the MLQ instrument, assuming that decision-making is part of transformational and transactional leadership. However, it is quite difficult to assume as similar forms of transformational leadership the cases of one leader that listen and involves the followers toward the mission to accomplish (participative decision-making) to another leader that defines alone the mission to accomplish and then communicates to the followers what they are all trying to accomplish (autocratic decision-making). In the similar way, for transactional leadership, it is difficult to assume as similar forms of leadership the cases of a leader that negotiates with followers the rewards to attribute if they succeed in accomplishing the mission regarding other leader that established alone the rewards and then communicates the decision to the followers. Thus, introducing decision-making in the comprehension of leadership may increase our knowledge of how leaders influences followers toward the established mission.

In sum, full range leadership theory and MLQ-5X represents important contributions for the study of leadership but they can be refined and improved. This is the main goal of this study by providing a comparison between the MLQ-5X and the Multidimensional Sport Leadership Scale (MSLS; Gomes & Resende, 2014). MSLS evaluates nine factors of leadership distributed for three styles of leadership (transformational, transactional, and decision-making). The development of MSLS subscales and items followed four criteria that helped to defined the nine factors of leadership: (a) Theoretical observation: leadership' factors correspond to specific behaviors that can be observed (and identified) when the leader relates to followers and teams; (b) Theoretical variance: leadership' factors includes different behaviors that share the same goal of leadership, and because of that they can be organized together; (c) Theoretical independence: each leadership' factors should be perceived equally by the leader and the followers and should

be perceived distinctly from others sets of leaders' behaviors; and (d) Theoretical impact: leadership' factors achieve "usefulness" when it is possible to establish, positive or negative, relations with subjective or objective measures of leadership efficacy (Gomes, 2019). Thus, transformational leadership includes five factors: (1) Vision: leader' ability to present an enthusiastic and optimistic vision of followers' future; (2) Inspiration: leader' positive expectations and behaviors directed toward promoting the success and continuous efforts of followers; (3) Instruction: leaders' actions focused on teaching and improving followers' technical skills; (4) Respect: leaders' tendency to consider the needs and personal and professional expectations of followers; and (e) Support: leader' personal concern regarding followers' well-being and interest in building positive relationships based on confidence. The development of items for each instrument had in common some type of transformational changing produced by the leader. For example, dimensions of Instruction and Support were formulated not only to capture task and relationship orientated leadership but also to capture how these aspects stimulated some type of changing on followers. For example, one item of Instruction refers "My ... explains me what should be done in order to improve my skills", meaning that the item not only evaluates the technical ability of the leader ("explains me what should be done...") but also the impact produced by this instruction on followers abilities ("...in order to improve my skills"). Transactional leadership includes two factors: (6) Positive Feedback: leader' reinforcement and recognition of good performance and efforts of followers; and (7) Negative Feedback: leader' punishment behaviors intended to manage followers' inadequate behaviors and performance. In the MSLS it is not included factors of Management-by-Exception as in MLQ-5X; some of these behaviors were included in the decision-making domain.

Decision-making leadership includes two factors: (8) Active Management: leader' behaviors of power management, by assuming the decisions in a more decentralized process (involving team members) or in a more centralized process (assuming the decisions alone); and (9) Passive Management: leader' avoidance or delay in taking responsibility for decision-making when it is necessary to solve important problems. In the MSLS, aspects of Passive Management-by-Exception and Laissez-faire are evaluated together in Passive Management; however, Active Management-by-Exception of MLQ-5X do not corresponds to Active Management of MSLS; the first is related to monitor and acting if deviations from rules and standards occur while Active Management is related to evaluate decision making as more centralized (low scores on the subscale) or more decentralized (higher scores on the subscale). Figure 1 summarizes the relations between MSLS and MLQ-

5X subscales, giving a better idea of correspondences between both instruments.

The MSLS has been most used in sports contexts to explain the coaches' leadership efficacy (Gomes & Resende, 2014; Mata & Gomes, 2013). However, there are no indications of the advantages of using this instrument instead of the MLQ-5X, and as consequence it is not possible to determine if the instrument improves the actual knowledge provided by the MLQ-5X on understanding leadership which remains a central topic of leadership studies (Hoch et al., 2018). In this study, we compared both instruments according the athletes' perspective of coaches' leadership. The choice for sports leadership was because it represents a fertile field for study coach-athletes relationships as determinant to explain the well-being and performance of athletes (Callow, Smith, Hardy, Arthur, & Hardy, 2009; Gomes, 2014; Stenling &

Multidimensional Sport Leadership Scale (MSLS)	Multifactorial Leadership Questionnaire (MLQ-5X)
<u>Transformational leadership</u>	<u>Transformational leadership</u>
Vision	Idealized influence: Attributed Idealized influence: Behavior
Inspiration	Inspirational motivation
Instruction	--
--	Intellectual stimulation
Individualization	Individualized consideration
Support	--
<u>Transactional leadership</u>	<u>Transactional leadership</u>
Positive feedback	Contingent reward
Negative feedback	--
<u>Decision-making leadership</u>	<u>Transactional leadership</u>
Active management	Management-by-exception: Active
Passive management	Management-by-exception: Passive
	<u>Nonleadership</u>
--	Laissez-faire

Fig. 1. - Correspondence between MSLS and MLQ-5X subscales.

Tafvelin, 2014). Besides, we selected for this study high competition athletes because the topic of leadership is more decisive for high performance and sport success. Although MLQ-5X was not first developed for sport contexts, there are positive indications of using the instrument for study coach-athlete relationships (see Rowold, 2006).

In sum, we conduct a validity analysis of the MSLS by comparing with the MLQ-5X. Specifically, we compared the psychometric properties of the MSLS and the MLQ-5X (goal 1), and to analyze the convergent and discriminant validity of both leadership instruments (goal 2). For the first goal, we formulated one hypothesis:

H1. The MSLS represents a better measure to evaluate the coaches' leadership than the MLQ-5X (H1) because it includes a set of factors more adjusted to evaluate transformational and transactional leadership and includes the domain of decision-making.

In order to consolidate the results and conclusions of H1 we formulated three associated hypotheses that constitutes the second goal of this study:

H2. The MSLS subscales would show convergent validity to constructs of their own instrument than to the counterpart subscales of MLQ-5X.

H3. The MLQ-5X subscales would show convergent validity to constructs of their own instrument than to the counterpart subscales of MSLS.

H4. The MSLS transformational and transactional constructs would show convergent validity to Satisfaction with Leadership.

Method

PARTICIPANTS

A total of 379 athletes (301 male, 79.4%; 78 female, 20.6%) from a variety of team sports (handball, 39%; football, 15%; volleyball, 12%; basketball, 9%; water polo, 9%; futsal, 9%; roller hockey, 7%) agreed to participate in this study. Their age ranged from 18 to 47 years-old ($M = 25.71$, $SD = 9.01$), and their sport experience from 1 to 31 years ($M = 13.36$, $SD = 7.34$). They were all adult athletes, competing in the national divisions of their sports (65% 1st division, 35% 2nd division). Regarding the athletes' relationship with their current coach, 57% of athletes worked with their coach for more than one year, and 36% of athletes won at least one regional, national, or international trophy in the past five years with their current coach.

PROCEDURE

Local teams and clubs were contacted, and athletes who agreed to participate in the study were asked to complete a questionnaire regarding their perceptions about their coach's behavior and leadership style, as well as their satisfaction with their sport activity. Athletes fulfilled an informed consent about the goals and voluntary participation in the study and then

completed the questionnaire before a training session, in a private room without the presence of their coach. All athletes had, at least, two months of working together with their current coach. The study received approval from the Ethics Committee of the first authors' institution.

INSTRUMENTS

MULTIDIMENSIONAL SPORT LEADERSHIP SCALE (MSLS; GOMES & RESENDE, 2014)

This 36-item instrument evaluates athletes' perceptions of leadership behaviors from their coaches, asking them to rate how frequently their coach displayed a specific behavior (*1 = never, 5 = always*). It includes nine subscales (four items per subscale) distributed by the constructs of transformational leadership, transactional leadership, and decision-making. Transformational leadership is composed by the following subscales: (1) Vision (e.g. "My coach defines a positive vision for my future"), (2) Inspiration (e.g. "My coach encourages me to work the best I can"), (3) Instruction (e.g. "My coach tells me what to do and how to do it, facilitating my progress in this sport"), (4) Individualization (e.g. "My coach respects me as a person and not as «just» one more team member"), and (5) Support (e.g. "My coach helps me when I have a personal problem"). The transactional construct includes two subscales: (6) Positive Feedback (e.g. "My coach demonstrates satisfaction when I perform well"), and (7) Negative Feedback (e.g. "My coach gets irritated with me when I don't do things as planned"). Finally, the decision-making construct encompasses two more subscales (8) Active Management (e.g. "My coach encourages me to give suggestions about what to do in training/competitions"), and (9) Passive Management (e.g. "My coach allows problems go on before doing something about it"). The subscales were computed by averaging the items.

MULTIFACTORIAL LEADERSHIP QUESTIONNAIRE (MLQ-5X; AVOLIO & BASS, 2004)

The MLQ-5X short version (45-items) was translated from English to Portuguese by authors of this study and then back translated by an English native speaker with knowledge in master degree in psychology (Behling & Law, 2000). The term "leader" was replaced by "coach" to make it more meaningful to athletes. For our study we excluded the three subscales of extra effort, satisfaction with the leader, and the effectiveness of the leader (total of 9 items) because they were not related with central dimensions of transformational leadership, transactional leadership, and nonleadership. Thus, MLQ-5X included 36 items responded in a 5-point scale (*0 = never, 4 = always*) divided by transformational leadership (five subscales), transactional leadership (three subscales) and laissez-faire leadership (one subscale), resulting in nine subscales (four items each). Specifically, the transformational leadership construct comprises the following subscales: (1) Idealized influence: Attributed (e.g. "Goes beyond self-interest for the good of the group"), (2) Idealized influence: Behavior (e.g. "Talks about his/her most important values and beliefs"), (3) Inspirational Motivation (e.g. "Talks optimistically about the future"), (4) Intellectual Stimulation (e.g. "Gets me to look at problems from many different angles"), and (5) Individualized Consideration (e.g. "Treats me as an individual rather than just as a member of a group"). The three transactional leadership subscales are (6) Contingent Reward (e.g. "Provides me with assistance in exchange for my efforts"), (7) Management-by-exception: Active (e.g. "Keeps track of all mistakes"), and (8) Management-by-exception: Passive (e.g. "Fails to interfere until problems become serious"). The final subscale represents the absence of leadership and is called (9) *Laissez-faire* (e.g. "Avoids getting involved when important issues arise"). Each subscale was computed by averaging the score of their items.

ATHLETE SATISFACTION QUESTIONNAIRE (ASQ; RIENER & CHELLADURAI, 1998).

The Portuguese version of the instrument was used in this study (Gomes & Resende, 2014), namely the four subscales related to Satisfaction with Leadership (19 items). Athletes were asked to rate their satisfaction (1 = *not satisfied at all*, 7 = *extremely satisfied*) with four aspect of the leadership of their coaches: (1) Ability strategy ("I am satisfied with the degree to which my abilities are used"), (2) Strategy ("I am satisfied with the tactics used during games"), (3) Personal treatment ("I am satisfied with the recognition I receive from my coach"), and (4) Training and Instruction ("I am satisfied with the training I receive from the coach during the season"). A Confirmatory Factor-Analysis revealed an acceptable fit level of this structure with a second order factor called "Satisfaction with Leadership", $\chi^2(143) = 478.11$, $p < .001$, $\chi^2/df = 3.34$, RMSEA = .079, 90% CI [.071; .087], p (RMSEA $\leq .05$) = .001, SRMR = .062, CFI = .949, NFI = .929; TLI = .939, IFI = .949. Thus, a global satisfaction score was computed by averaging participants' responses ($\alpha = .96$ for this study).

DATA ANALYSIS

We used a two-step strategy in order to analyze the psychometric properties of both leadership instruments and to compare their discriminant validity. The analyses were conducted using SPSS (v. 25) and AMOS (v. 24; SPSS Inc. Chicago, IL. USA).

The first step consisted on assessing the construct validity of each instrument (MSLS and MLQ-5X) by using three analytical approaches: (1) confirmatory factor analysis, (2) instrument reliability, and (3) convergent and discriminant validity. The confirmatory factor analysis (CFA) was conducted to test the latent structure of the leadership instruments (MSLS and MLQ-5X). Structural models were tested through maximum likelihood (ML) estimation methods. To estimate models fit, and know how well theoretical models were able to reproduce the correlational structure of the observed variables, we used the following indicators: (a) the chi-square goodness-of-fit statistics (χ^2), and the absolute index χ^2/df (Arbuckle, 2008); (b) the Steiger-Lind Root Mean Square Error of Approximation (RMSEA), its 90% confidence interval and p -value for H0: RMSEA $\leq .05$ (Steiger, 1990); (c) the root-mean-square residual (SRMR; Marôco, 2014); (d) three relative indices, the Comparative Fit Index (CFI) (Bentler, 2007), the Tucker-Lewis Index (TLI) (Bentler & Bonett, 1980; Tucker & Lewis, 1973) and the Incremental Fit Index (IFI) (Bollen, 1989); and, (e) the Akaike Information Criterion (AIC) (Arbuckle, 2008). The cutoff measures followed the most acceptable criteria described in the literature (Kline, 2016). Values of RMSEA in the range [.05, .08] constituted an acceptable fit, and inferior to .05 indicated an excellent fit (Arbuckle, 2008). Values of SRMR of .05 and .10 are considered good and adequate fits (Marôco, 2014); Values of CFI, TLI and IFI in the range [.90, .95] indicated a good adjustment, and values higher than .95 an excellent adjustment (Bollen, 1989). With Respect to AIC, small values compared to other competing models, were indicative of a good-fitting parsimonious model (Tabachnick & Fidell, 2013). The chi-square goodness-of-fit statistics, testing for the significance of the discrepancy function, indicated the best fit of the global model to data, for a lower statistical test value, and probability higher than .05 (Kline, 2016). Instruments' reliability was assessed through two different measures: the reliability of each item was assessed through regression weights (for all items $\lambda \geq .50$, $R^2 \geq .25$; (Hair, Black, Babin, & Anderson, 2014), whilst construct reliability was evaluated through the Cronbach's alpha measure of internal consistency and the composite reliability measure (CR) (Hair et al., 2014). A construct reliability was considered acceptable if the values of Cronbach's alpha and CR were equal or greater than .70 (Hair et al., 2014; Marôco, 2014). Finally, the assessment of instruments' convergent and discriminant validity followed Fornell and Larcker's (1981) procedure. Convergent validity was estimated through the average variance extracted (AVE) and the CR. To assume convergent validity, AVE should be equal or greater than .50 and lower than CR – this would ensure that the variance explained by the construct is higher than the measurement error, and larger than the

cross-loadings (Hair et al., 2014). Discriminant validity was assessed by analyzing whether the average variance extracted (AVE) was greater than the maximum shared squared variance (MSV) or the average shared squared variance (ASV). According to this method (cf. Fornell & Larcker, 1981), the shared variance between two subscales within the same instrument (R^2_{ij}) should be inferior to the average of the AVEs for these, as well as lower than the AVE of each subscale (AVE_i and $AVE_j \geq R^2_{ij}$).

The second step focused on comparing the two instruments against each other. To do so, we considered two different approaches: (1) convergent and divergent validity between the two instruments of leadership (i.e., MSLS and MLQ-5X); (2) predictive validity of the MSLS and MLQ-5X, regarding athletes' Satisfaction with Leadership (ASQ). Figure 2 presents the methodological schema and rational used for discriminant validity analysis between the leadership measures; we retained three statistical indicators: (1) the correlation value between the subscale and its own construct (r_1); (2) the correlation value between the subscale and its counterpart construct from the other instrument (r_2); and, (3) the correlation value between the constructs of both instruments (r_3), which was target of control in statistical analysis. We used the Z-statistics, as proposed by Meng, Rosenthal, and Rubin (1992), to test if a Construct A would be stronger associated with a Construct B (convergent validity) than with a Construct C (divergent validity). In this method, we tested for significant differences between the correlation coefficients, $r_{1/AB}$ and $r_{2/AC}$, while controlling for $r_{3/BC}$, and rejected the null hypothesis of equality when p value < .05 (Meng et al., 1992).

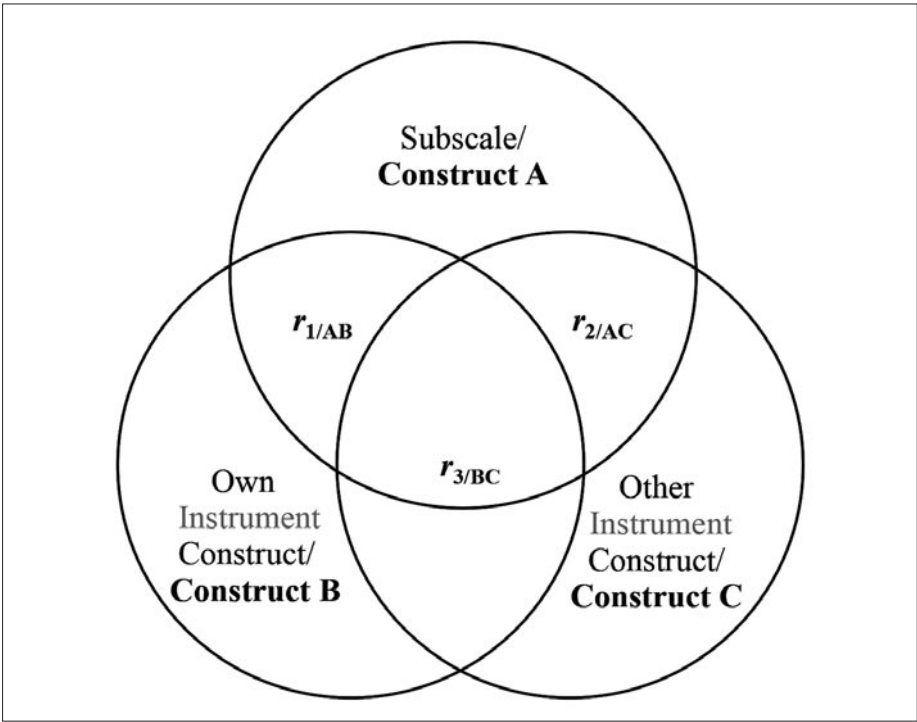


Fig. 2. - Graphic representation of the Schema used in Convergent Validity Analysis.

Results

DESCRIPTIVE STATISTICS AND CORRELATIONS

Table I summarizes descriptive statistics and correlations among MSLS and MLQ-5X subscales scores and ASQ global score.

As shown in Table 1, Satisfaction with leadership (ASQ) is positively correlated with all subscales of leadership instruments, with exception of Passive management (MSLS), Management by exception passive (MLQ-5X), and Laissez-faire (MLQ-5X) where correlations were negative. There was no correlation between Satisfaction with leadership and Negative feedback (MSLS).

Regarding Management by exception passive (MLQ-5X) and Laissez-faire (MLQ-5X), these subscales were positively correlated with Passive management (MSLS) and negatively correlated with the remaining MSLS subscales, with exception of Negative feedback (only positively correlated with Management by exception passive) and Active Management (no significant correlation). The remaining MLQ-5X subscales were all negatively related with Passive management (MSLS), and positively correlated with the remaining MSLS subscales, with exception of Negative feedback (MSLS; which was only positively correlated with Intellectual stimulation and Management by exception active).

PSYCHOMETRIC PROPERTIES OF LEADERSHIP INSTRUMENTS

Latent Leadership Structure. Regarding the latent structure for the leadership instruments - MSLS and MLQ-5X - results from the CFA showed that, the models tested fitted the data well. Specifically, for the 9-factor model of the MSLS, the following tests of significance and goodness-of-fit measures were obtained: $\chi^2(558) = 1127.59$, $p < .001$, $\chi^2/df = 2.02$, RMSEA = .052, 90% CI [.048; .056], p (RMSEA $\leq .05$) = .226, SRMR = .057, CFI = .928, TLI = .919, IFI = .929, AIC = 1343.59, BIC = 1768.84. The fit level for the 9-factor structure of the MLQ-5X was attained by the following measures: $\chi^2(558) = 1206.93$, $p < .001$, $\chi^2/df = 2.16$, RMSEA = .055, 90% CI [.051; .060], p (RMSEA $\leq .05$) = .018, SRMR = .059, CFI = .912, TLI = .900, IFI = .912, AIC = 1422.93, BIC = 1848.18. Therefore, considering the goodness-of-fit measures shown by the two leadership instruments, the 9-factor model of the MSLS was considered the most favorable, in comparison to the 9-factor structure of the MLQ-5X. This is supported by AIC and BIC values, which are lower in the MSLS model when compared to the MLQ-5X.

TABLE I
Means, Standard Deviations And Correlations Between The Variables Under Investigation (N = 301).

MSLS	M (SD)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Vision	3.47 (1.00)	—																		
2. Inspiration	3.85 (0.86)	.71**	—																	
3. Instruction	3.91 (0.81)	.50**	.75**	—																
4. Individualization	4.05 (0.76)	.56**	.60**	.49**	—															
5. Support	3.07 (0.97)	.56**	.54**	.41**	.59**	—														
6. Positive Feedback	3.51 (0.86)	.60**	.64**	.47**	.57**	.53**	—													
7. Negative Feedback	2.68 (0.90)	.03	.08	.11*	.22**	.07	-.07	—												
8. Active Management	2.80 (0.95)	.47**	.47**	.33**	.44**	.57**	.50**	-.03	—											
9. Passive management	2.12 (0.87)	-.27**	-.47**	-.53**	-.30**	-.13**	-.23**	.10	-.03	—										
MIQ-5X																				
10. Idealized influence: Attributed	2.76 (0.81)	.63**	.75**	.64**	.59**	.55**	.61**	.02	.48**	-.43**	—									
11. Idealized influence: Behavior	2.87 (0.72)	.49**	.66**	.66**	.57**	.50**	.50**	.02	.36**	-.43**	.72**	—								
12. Inspirational motivation	2.99 (0.73)	.62**	.71**	.63**	.58**	.53**	.58**	-.04	.42**	-.41**	.74**	.75**	—							
13. Intellectual Stimulation	2.50 (0.79)	.56**	.73**	.72**	.49**	.56**	.51**	.12*	.47**	-.42**	.75**	.70**	.71**	—						
14. Individualized Consideration	2.70 (0.80)	.62**	.75**	.67**	.61**	.62**	.59**	.05	.52**	-.38**	.77**	.68**	.71**	.76**	—					
15. Contingent Reward	2.68 (0.81)	.64**	.72**	.59**	.60**	.55**	.72**	-.01	.46**	-.35**	.76**	.69**	.70**	.73**	.73**	—				
16. Management-by-exception Active	2.63 (0.70)	.41**	.63**	.64**	.34**	.37**	.37**	.14**	.31**	-.39**	.65**	.64**	.61**	.66**	.58**	.58**	—			
17. Management-by-exception: Passive	1.61 (0.82)	-.19**	-.35**	-.35**	-.26**	-.12*	-.17**	.10*	-.07	.65**	.32**	-.32**	-.32**	-.33**	-.31**	-.23**	-.23**	—		
18. Laissez-faire	0.99 (0.83)	-.24**	-.41**	-.42**	-.25**	-.13*	-.16**	.07	.01	.67**	-.37**	-.41**	-.39**	-.38**	-.36**	-.29**	-.34**	.71**	—	
ASQ																				
19. Satisfaction with Leadership	4.87 (1.11)	.65**	.78**	.66**	.65**	.56**	.65**	-.07	.49**	-.43**	.74**	.65**	.70**	.71**	.76**	.74**	.58**	-.35**	-.39**	—

Note. * $p < .05$. ** $p \leq .01$.

Subscales' Reliability. For the reliability of the individual items, CFA regression weights for the MLQ-5X indicated that all of the items presented a $\lambda \geq .50$ ($R^2 \geq .25$), ranging from .51 to .83. The exceptions were item four, concerning Active Management-by-Exception subscale, with a $\lambda = .17$ and, item 17, concerning Passive Management-by-Exception subscale, with a $\lambda = .20$, showing a reduced contribution for the leadership constructs they represent¹. Regarding CFA regression weights for the MSLS, all of the items presented a $\lambda \geq .50$, ranging from .54 to .91. With no exception, all MSLS items showed a significant contribution for the constructs they represent. Regarding construct reliability, and the composite reliability (CR) criteria, the MSLS and MLQ-5X did not present problems in this domain, showing a good reliability of the leadership subscales ($CR \geq .70$) (Table 2). Even though the Cronbach's alpha criteria of two of the MLQ-5X subscales (Management-by-Exception Active and Management-by-Exception Passive) assumed problems of internal consistency, their values were near the acceptable (i.e., $\alpha = .687$ and $\alpha = .696$, for this study). On the contrary, all MSLS subscales showed good levels of internal consistency (i.e., $.761 \leq \alpha \leq .921$, for this study) (Table II).

SUBSCALES' CONVERGENT AND DISCRIMINANT VALIDITY

Convergent validity analysis for the MLQ-5X showed that, besides the values given by the CFA being quite acceptable and the AVE for each MLQ-5X-subscale being lower than the CR measure, the AVE presented by five subscales was inferior to .50 (see Table II). To this point, we should notice that when the AVE is lower than .50 it is difficult to differentiate between variance due to construct and variance due to error of measurement. In addition, it is important to remember that items with low regression weights and cross-loading problems (e.g., items 4 and 17) influenced those results. Even though, for the purpose of this study, we kept the items because of their theoretical value, and since other reliability measures criteria were achieved

¹ As the cultural adjustments were ensured during the translation process (specifically, we used a back-translation method), and due to the wide use of MLQ-5X, we decided to keep all items of both scales in order to ensure a fair comparison. However, as informative note, we analyzed MLQ-5X without items 4 and 17: $\chi^2(491) = 1052.458$ $p < .001$, $\chi^2/df = 2.14$, RMSEA = .055, 90% CI [.050; .060], p (RMSEA $\leq .05$) = .037, SRMR = .053, CFI = .922, TLI = .911, IFI = .923, AIC = 1260.46, BIC = 1669.96. Management-by-exception: Passive (without item 17): Cronbach's $\alpha = .766$; CR = .779; AVE = .546; MSV = .872; Management-by-exception: Active (without item 4): Cronbach's $\alpha = .772$; CR = .770; AVE = .527; MSV = .958.

TABLE II
Construct Reliability And Validity For The MSLS And MLQ-5X Subscales (N = 301).

Leadership Instrument	Cronbach's	CR	AVE	MSV
MSLS				
Vision	.921	.922	.748	.621
Inspiration	.876	.880	.648	.752
Instruction	.831	.835	.559	.752
Individualization	.824	.834	.557	.487
Support	.845	.848	.584	.486
Positive Feedback	.850	.858	.648	.489
Negative Feedback	.813	.820	.539	.074
Active Management	.830	.831	.552	.404
Passive Management	.761	.767	.456	.432
MLQ-5X				
Idealized influence: Attributed	.765	.770	.461	.970
Idealized influence: Behavior	.755	.771	.461	.899
Inspirational motivation	.848	.850	.586	.870
Intellectual Stimulation	.819	.821	.537	.966
Individualized Consideration	.749	.756	.445	.970
Contingent Reward	.814	.814	.523	.901
Management-by-exception: Active	.687	.703	.409	.939
Management-by-exception: Passive	.696	.713	.422	.870
Laissez-faire	.811	.811	.518	.870

Note. CR = Composite Reliability Measure. AVE = Average Variance Extracted. MSV = Maximum Shared Squared Variance.

(e.g., $CR \geq .70$). Concerning MSLS, with the exception of Passive Management subscale ($AVE = .456$) all the criteria for convergent validity were fulfilled for the other subscales ($AVE \geq .50$ and $AVE \geq CR$) (see Table 2). Thus, eight MSLS subscales represented well the respective theoretical latent construct while for MLQ-5X only four subscales presented criteria for construct convergent validity.

Discriminant validity analysis tested the extent to which a subscale is truly distinct from other subscales, in terms of how much it correlates with other subscales and how distinctly measured variables that represent only the respective single construct of the subscale. Results for the MLQ-5X revealed problems for all leadership subscales, since the AVE was lower than their own MSV. Regarding MSLS, only two leadership subscales (Inspiration and Instruction) did not meet the criteria for discriminant validity, since the $AVE_{Subscale} < MSV_{Subscale}$ (see Table 2). Accordingly, seven MSLS subscales met conditions of discriminant validity, showing that the concepts they represent were distinct, while none of the MLQ-5X subscales exhibited criteria for construct discriminant validity.

In sum, results from CFA, subscales' reliability, convergent and discriminant analysis confirmed the central hypothesis of this study (H1). Based on

AVE, MSV and reliability indicators, as well as on the model fit indexes, the MSLS presents more satisfactory psychometric properties than MLQ-5X.

DISCRIMINANT VALIDITY BETWEEN LEADERSHIP MEASURES

In this step of data analysis, we analyzed convergent and divergent validity between the MSLS and MLQ-5X, introducing in the analysis the Athlete Satisfaction Questionnaire (ASQ). To do so, first, we tested for convergent validity of MSLS and MLQ-5X subscales with their leadership constructs of Transformational Leadership (TFL), Transactional Leadership (TSL), Decision-Making (DM), and Laissez Faire (LF). Second, we tested for convergent validity of MSLS and MLQ-5X constructs of transformational and transactional leadership with Satisfaction with Leadership (ASQ). Convergent validity of MSLS and MLQ-5X subscales with leadership constructs followed two steps of analysis. First, we tested for convergent validity of MSLS subscales with their own constructs and MLQ-5X counterparts. Second, we examined for convergent validity of MLQ-5X subscales with their own constructs and MSLS counterparts.

Comparing MSLS and MLQ-5X Subscales with Leadership Constructs. Regarding the first step of analysis, as displayed in Table 3, the MSLS transactional and decision-making subscales presented a significantly stronger association with their own constructs (i.e., from the MSLS), than with their counterparts from the MLQ-5X. For the transformational subscales, the exception goes for Instruction, which was as strongly related to MSLS-TFL as with MLQ-TFL constructs ($p = .24$). Therefore, except for one MSLS transformational subscale, data confirmed convergent validity for all subscales regarding the constructs of their own instrument, supporting hypothesis H2. The second step of analysis, showed that all MLQ-5X transformational and transactional subscales presented a significantly stronger association to their own instrument constructs, than to their counterparts from the MSLS ($p < .001$). Therefore, findings confirmed hypothesis H3, showing convergent validity for the MLQ-5X subscales regarding the constructs of their own instrument (see Table III).

Comparing MSLS and MLQ-5X Constructs with QSA. The transformational constructs from both MSLS and MLQ-5X presented a positive and equally significant stronger correlation with athletes' Satisfaction with Leadership (ASQ) ($\Delta r = 0.01$, $Z = 0.66$, $p = .508$, 95%CI [-0.06, 0.11]), thus discarding discriminant validity between those constructs. Regarding transactional leadership, the MLQ-5X transactional construct presented a positive

TABLE II
Z-Statistics For MSLS And MLQ-5X Subscales And Constructs Of Transformational Leadership,
Transactional Leadership And Decision-Making (N = 301).

SUBSCALES	CONSTRUCTS		Δr	Z	p	95% CI
	Pearson's r	Correlation				
	MSLS-TFL	MLQ-TFL				
MSLS-TFL	—	.86***				
Vision	.83**	.66**	0.17	10.05	<.001	[0.31, 0.44]
Inspiration	.89**	.82**	0.07	5.59	<.001	[0.17, 0.34]
Instruction	.77**	.75**	0.02	1.19	.24	[-0.03, 0.12]
Individualization	.77**	.64**	0.13	7.04	<.001	[0.19, 0.32]
Support	.78**	.62**	0.16	8.63	<.001	[0.24, 0.37]
MLQ-TFL	.86***	—				
Idealized influence: Attributed	.78**	.90**	-0.12	-9.26	<.001	[-0.48, -0.32]
Idealized influence: Behavior	.71**	.87**	-0.16	-10.58	<.001	[-0.48, -0.35]
Inspirational motivation	.76**	.88**	-0.12	-8.57	<.001	[-0.44, -0.28]
Intellectual Stimulation	.76**	.89**	-0.13	-9.51	<.001	[-0.47, -0.33]
Individualized Consideration	.81**	.89**	-0.08	-6.28	<.001	[-0.37, -0.20]
	MSLS-TSL	MLQ-TSL				
MSLS-TSL	—	.46***				
Positive Feedback	.67**	.51**	0.16	4.03	<.001	[0.13, 0.35]
Negative Feedback	.70**	.12*	0.58	12.64	<.001	[0.56, 0.70]
MLQ-TSL	.46***	—				
Contingent Reward	.51**	.74**	-0.23	-6.18	<.001	[-0.47, 0.26]
Management-by-exception: Active	.37**	.71**	-0.34	-8.21	<.001	[-0.55, -0.36]
Management-by-exception: Passive	-.04	.35**	-0.39	-7.39	<.001	[-0.47, -0.29]
	MSLS-DM	MLQ-TSL				
MSLS-DM	—	.28***				
Active Management	.73**	.39**	0.34	7.51	<.001	[0.36, 0.57]
Passive Management	.67**	-.01	0.68	12.34	<.001	[0.60, 0.74]

Note. CI = Confidence Interval. TFL = Transformational Leadership. TSL = Transactional Leadership. DM = Decision-making. Z-statistics includes the correlation between the two constructs of transformational leadership (MSLS-TFL and MLQ-TFL), transactional leadership (MSLS-TSL and MLQ-TSL) and decision-making (MSLS-DM and MLQ-TSL). * $p < .05$. ** $p \leq .01$. *** $p \leq .001$.

and significantly stronger correlation with athletes' satisfaction, than the MSLS homologous construct did ($\Delta r = -.10$, $Z = -2.18$, $p = .029$, 95%CI [-0.24, -0.01]). Therefore, results of athletes' Satisfaction with Leadership showed discriminant validity only for the transactional constructs, but not in the expected way, rejecting partially hypothesis H4 of convergent validity for MSLS transformational and transactional constructs.

Discussion

In this study, we conducted a validity analysis of the Multidimensional Sport Leadership Scale (MSLS), by establishing a comparison to the well-known leadership instrument, the Multifactorial Leadership Questionnaire (MLQ-5X). Specifically, we formulated two main goals: (1) compare the psychometric properties of the two leadership measures, formulating the central hypothesis of the study (H1); and, (2) analyze discriminant validity between the two instruments, formulating three additional hypothesis related to hypothesis 1 (H2, H3, and H4).

In the first hypothesis (H1), we expected the MSLS would present better psychometric properties than the MLQ-5X. Findings from CFA, instruments' reliability, and construct convergent and discriminant validity confirmed the hypothesis. Specifically, the construct validity (CFA for both instruments) revealed that the 9-factor model of the MSLS presented the most favorable goodness-of-fit measures. Also, instruments' reliability of the MSLS showed to be the most consistent measure in terms of the individual items' reliability and construct reliability. When analyzing Cronbach alpha's values, the MLQ-5X presented problems of internal consistency in two subscales (i.e., Management-by-Exception Active and Management-by-Exception Passive), even though the same problems were not found when analyzing CR values. Additionally, we found problems of convergent validity in five MLQ-5X subscales, turning difficult to distinguish between variance due to construct and variance due to error of measurement, and problems of discriminant validity in all MLQ-5X subscales. These problems are according with other studies that reported lack of construct validity and discriminant validity among constructs of the instrument (Antonakis et al., 2003; Bycio, Hackett, & Allen, 1995; Carless, 1998; Hinkin & Schriesheim, 2008; Tejada et al., 2001), resulting in a need of better clarify how transformational and even transactional leadership are evaluated by the MLQ-5X. For the MSLS, only Passive Management did not fulfilled all the criteria for convergent validity but had acceptable discriminant validity besides acceptable levels of reliability. One possible reason for this result may be related to the fact that Passive Management does not represents a typical behavior of sports coaches, meaning that they can be perceived by athletes as more active in trainings and competitions. In fact, Passive Management was the factor with lower mean of all nine MSLS factors ($M = 2.12$) but this value was substantially superior than Management-by-exception: Passive ($M = 1.61$) and Laissez-faire of MLQ-5X with residual value (0.99), meaning that it can be more sensible to capture situations where coaches fail to assume responsibility

when athletes expect an action from them. Other aspect that should be considered is that Inspiration and Instruction did not achieved all the criteria for discriminant validity and had higher value of correlation ($r = .75$); this may reflect the possibility that when coaches instruct athletes to improve their skills they are also perceived as inspirational; nevertheless, Inspiration and Instruction achieved acceptable levels of convergent validity besides acceptable levels of reliability. All in all, results from H1 revealed a better factorial structure for MSLS when comparing with MLQ-5X. Results of construct validity of MSLS are slightly better than the ones found by Gomes and Resende (2014) which may indicate structural stability of the instrument between both periods of data collection. It is also interesting to note that even the MLQ-5X achieved better RMSEA than the study of adaptation the instrument for sports context done by Rowold (2006); however, our data extends findings of Rowold and even Avolio and Bass (2004) by providing construct convergent and discriminant validity results that were not so positive for MLQ-5X.

In the second hypothesis (H2), we expected the MSLS subscales would show convergent validity to constructs of their own instrument. Findings confirmed the hypothesis, showing that with the exception of Instruction (that seems related to both constructs of transformational leadership of MSLS and MLQ-5X) all of the other transformational, transactional, and decision-making subscales presented convergent validity to their MSLS constructs. In the third hypothesis (H3), we expected the MLQ-5X subscales would show convergent validity with constructs of their own instrument. Results confirmed the hypothesis, since all MLQ-5X transformational and transactional subscales exhibited convergent validity with the MLQ-5X constructs they represent. Results from both hypotheses are quite interesting because although both instruments evaluates similar constructs of transformational and transactional leadership, when analyzing their specific factorial structures we can conclude that they are distinct from each other. Also important, the introduction of decision-making is sustained by convergent validity results that integrates both factors of Active and Passive Management on the original MSLS instrument by comparison with MLQ-5X instrument. Thus, instruments converge to their respective factorial structure and this convergence suggests differences on both instruments on how constructs are evaluated, with the single exception of Instruction that did not achieve statistically significant results on convergent validity but it had higher values of relation with original MSLS instrument, suggesting that it “belongs” more to MSLS instrument than to MLQ-5X. These results indicated that we are in presence of two related but distinct instruments, meaning that they can con-

tribute, on their own, to understand the influence produced by leaders on individuals and organizations. These results, in conjunction with better results of reliability, construct convergent, and discriminant validity of MSLS when compared with MLQ-5X (H1), suggests that MSLS allow a more integrative analysis of leadership, which may have the advantage of diminish the construct redundancy on the definition and evaluation of leadership (Hoch et al., 2018).

In the fourth hypothesis (H4), we expected the MSLS transformational and transactional constructs would show convergent validity with Satisfaction with Leadership. The hypothesis was partially confirmed for the transformational factors of MSLS and MLQ-5X instruments because they assumed convergent validity in explaining athletes' Satisfaction with Leadership. However, this was not the case for transactional factors of MSLS and MLQ-5X; MLQ-5X transactional constructs assumed stronger correlation with athletes' satisfaction than the MSLS homologous construct. One potential reason for this result is that MLQ-5X add three subscales contributing to convergent validity with Satisfaction with Leadership while MSLS had only two. However, as verified on H1, the transactional leadership of MLQ-5X had problems of construct validity and discriminant validity, which seems confirmed by other studies (Hinkin & Schriesheim, 2008; Tejada et al., 2001). Besides, the fact that transformational leadership of both instruments did not differed on convergent validity regarding Satisfaction with Leadership is worth of note. In fact, research on the augmentation effect of transformational leadership regarding transactional leadership (Bass, 1985) indicates that transformational leadership adds unique variance beyond that of transactional leadership for predicting subjective and objective indicators of leadership efficacy (Birasnav, 2014; Gomes & Resende, 2014; Rowold, 2006; Zwingmann et al., 2014). For example, meta-analysis of Lowe, Kroeck, and Sivasubramaniam (1996) as well as Judge and Piccolo (2004) demonstrates that correlation of transactional leadership and outcomes of leadership are positive, but to a lesser degree than with transformational leadership. There is no doubt that transformational leadership can augment our knowledge of effects produced by leaders in a wide range of contexts, including profit and nonprofit organizations (Lowe et al., 1996), public sector (Rafferty & Griffin, 2004), educational organizations (Beauchamp et al., 2010), military contexts (Hardy et al., 2010), religious contexts (Rowold, 2008), and sports (Rowold, 2006).

In conclusion, actual knowledge about how to measure transformational leadership and other related constructs still needs to be improved (Hoch et al., 2018). Authors have been following distinct research options in order to

provide indications about alternative forms of evaluating leadership, as is the case of developing measures for specific contexts of leadership (Beauchamp et al., 2010), developing new measure of transformational leadership (Rafferty & Griffin, 2004), or comparing different measures of leadership (Rowold & Heinitz, (2007)). Our study aligns this last option by comparing two measures of leadership. Our data provided encouraging results for the MSLS which refines the way transformational and transactional leadership are evaluated and integrates the domain of decision-making. Future research may continue validation efforts of the instrument, by testing the factorial structure (it is still important to improve some convergent validity indicators), the stability of the instrument (our study had the cross-sectional limitation), and by extending data to other leadership contexts besides sports.

REFERENCES

- Antonakis, J., Avolio, B. J., & Sivasubramaniam, N. (2003). Context and leadership: An examination of the nine-factor full-range leadership theory using the Multifactor Leadership Questionnaire. *The Leadership Quarterly*, 14(3), 261–295.
- Antonakis, J., Bastardoz, N., Jacquart, P., & Shamir, B. (2016). Charisma: An ill-defined and ill-measured gift. *Annual Review of Organizational Psychology and Organizational Behavior*, 3, 293–319.
- Arbuckle, J. (2008). *AMOS 17 Users' Guide*. Chicago, IL: SPSS.
- Aryee, S., Chen, Z. X., Sun, L. Y., & Debrah, Y. A. (2007). Antecedents and outcomes of abusive supervision: Test of a trickle-down model. *Journal of Applied Psychology*, 92, 191–201.
- Avolio, B. J., & Bass, B. M. (1995). Individual consideration viewed at multiples levels of analysis: A multi-level framework for examining the diffusion of transformational leadership. *The Leadership Quarterly*, 6(2), 199–218.
- Avolio, B. J., & Bass, B. M. (2004). *Multifactor leadership questionnaire. Manual and sampler set* (3rd ed). Redwood City: Mind Garden, Inc.
- Banks, G. C., Gooty, J., Ross, R. L., Williams, C. E., & Harrington, N. T. (2018). Construct redundancy in leader behaviors: A review and agenda for the future. *The Leadership Quarterly*, 29(1), 236–251, doi: <https://doi.org/10.1016/j.leaqua.2017.12.005>.
- Bass, B. M. (1985). *Leadership and performance beyond expectations*. New York: Free Press.
- Bass, B. M. (1995). Theory of transformational leadership redux. *Leadership Quarterly*, 6(4), 463–478.
- Bass, B. M. (1998). *Transformational leadership: Individual, military and educational impact*. Mahwah, NJ: Erlbaum.
- Bass, B. M. (2008). *Bass handbook of leadership: Theory, research & managerial applications* (4rd Ed.). New York: The Free Press.
- Bass, B. M., & Avolio, B. J. (1991). *Multifactor Leadership Questionnaire*. Palo Alto, CA: Consulting Psychologists Press.
- Bass, B. M., & Avolio, B. J. (2000). *MLQ Multifactor Leadership Questionnaire*. Redwood City, CA: Mind Garden.
- Beauchamp, M. R., Barling, J., Li, Z., Morton, K. L., Keith, S. E., & Zumbo, B. D. (2010). Development and psychometric properties of the transformational teaching questionnaire. *Journal of Health Psychology*, 15(8), 1123–1134.
- Behling, O., & Law, K. S. (2000). *Translating questionnaires and other research instruments: Problems and solutions*. Thousand Oaks, CA: Sage.

- Bentler, P. M. (2007). On tests and indices for evaluating structural models. *Personality & Individual Differences*, 42(5), 825-829. doi:10.1016/j.paid.2006.09.024
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin*, 88(1), 588-606.
- Birasnav, M. (2014). Knowledge management and organizational performance in the service industry: The role of transformational leadership beyond the effects of transactional leadership. *Journal of Business Research*, 67(8), 1622-1629. doi: 10.1016/j.jbusres.2013.09.006
- Blanchard, K. (2007). *Leading at a higher lever: Blanchard on leadership and creating high performance organizations*. Upper Saddle River, NJ: Pearson-Prentice Hall.
- Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research*, 17(3), 303.
- Bycio, P., Hackett, R. D., & Allen, J. S. (1995). Further assessments of Bass's (1985) conceptualization of transactional and transformational leadership. *Journal of Applied Psychology*, 80(4), 468-478.
- Callow, N., Smith, M. J., Hardy, L., Arthur, C. A., & Hardy, J. (2009). Measurement of transformational leadership and its relationship with team cohesion and performance level. *Journal of Applied Sport Psychology*, 21(4), 395-412.
- Carless, S. A. (1998). Assessing the discriminant validity of transformational leadership behaviour as measured by the MLQ. *Journal of Occupational and Organizational Psychology*, 71(4), 353-358.
- Cheung G & Rensvold R (2002). Evaluating goodness-of-indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233-255. doi:10.1207/S15328007SEM0902_5
- DeRue, D. S., Nahrgang, J. D., Wellman, N., & Humphrey, S. E. (2011). Trait and behavioral theories of leadership: An integration and meta-analytic test of their relative validity. *Personnel Psychology*, 64(1), 7-52.
- Fornell, C., & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39-50.
- Gomes, A. R. (2014). Transformational leadership: Theory, research, and application to sports. In C. Mohiyeddini (Ed.), *Contemporary topics and trends in the psychology of sports* (pp. 53-114). New York: Nova Science Publishers.
- Gomes, A. R. (2019). Coaching efficacy: The Leadership Effectiveness Model. *Submitted for publication*.
- Gomes, A. R., & Resende, R. (2014). Assessing leadership styles of coaches and testing the augmentation effect in sport. In C. Mohiyeddini (Ed.), *Contemporary topics and trends in the psychology of sports* (pp. 115-137). New York: Nova Science Publishers.
- Goodwin, V. L., Wofford, J. C., & Whittington, J. L. (2001). A theoretical and empirical extension to the transformational leadership construct. *Journal of Organizational Behavior*, 22(7), 759-774.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis* (7th ed.). England: Pearson Education Limited.
- Hardy, L., Arthur, C.A., Jones, G., Shariff, A., Munnoch, K., Isaacs, I., et al. (2010). The relationship between transformational leadership behaviors, psychological, and training outcomes in elite military recruits. *The Leadership Quarterly*, 21(1), 20-32.
- Hinkin, T. R., & Schriesheim, C. A. (2008). Atheoretical and empirical examination of the transactional and non-leadership dimensions of the Multifactor Leadership Questionnaire (MLQ). *The Leadership Quarterly*, 19, 501-513.
- Hoch, J. E., Bommer, W. H., Dulebohn, J. H., & Wu, D. (2018). Do ethical, authentic, and servant leadership explain variance above and beyond transformational leadership? A meta-analysis. *Journal of Management*, 44(2), pp. 501-529. doi: https://doi.org/10.1177/0149206316665461
- Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology*, 89(5), 755-768.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York: Guilford Press.
- Lorinkova, N. M., Pearsall, M. J., & Sims Jr, H. P. (2013). Examining the differential longitudinal performance of directive versus empowering leadership in teams. *Academy of Management Journal*, 56(2), 573-596.

- Lowe, K. B., Kroeck, K. G., & Sivasubramaniam, N. (1996). Effectiveness correlates of transformational and transactional leadership: A meta-analytic review of the MLQ literature. *Leadership Quarterly*, 7(3), 385-425.
- Marôco, J. (2014). *Análise estatística com o SPSS Statistics* (6th ed.). [Statistical analysis with SPSS Statistics]. Pêro Pinheiro: ReportNumber, Lda.
- Mata, R. T., & Gomes, A. R. (2013). Winning or not winning: The influence on coach-athlete relationships and goal achievement. *Journal of Human Sport and Exercise*, 8(4), 986-995.
- Meng, X., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, 111(1), 172-175.
- Oc, B. (2018). Contextual leadership: A systematic review of how contextual factors shape leadership and its outcomes. *The Leadership Quarterly*, 29(1), 218-235.
- Pelletier, K. L. (2010). Leader toxicity: An empirical investigation of toxic behavior and rhetoric. *Leadership*, 6(4), 373-389. doi: <https://doi.org/10.1177/1742715010379308>
- Rafferty, A.E., & Griffin, M.A. (2004). Dimensions of transformational leadership: Conceptual and empirical extensions. *The Leadership Quarterly*, 15, 329-354.
- Riemer, H. A., & Chelladurai, P. (1998). Development of the athlete satisfaction questionnaire (ASQ). *Journal of Sport & Exercise Psychology*, 20(2), 127-156.
- Rowold, J. (2006). Transformational and transactional leadership in martial arts. *Journal of Applied Sport Psychology*, 18(4), 312-325.
- Rowold, J. (2008). Effects of transactional and transformational leadership of pastors. *Pastoral Psychology*, 56(4), 403-411. doi: 10.1007/s11089-008-0121-6
- Rowold, J., & Heinritz, K. (2007). Transformational and charismatic leadership: Assessing the convergent, divergent and criterion validity of the MLQ and the CKS. *The Leadership Quarterly*, 18(2), 121-133.
- Sosik, J. J., & Jung, D. (2018). *Full range leadership development* (2nd Edition). New York, NY: Routledge.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25(2), 173-180. doi: 10.1207/s15327906mbr2502_4
- Stenling, A., & Tafvelin, S. (2014). Transformational leadership and well-being in sports: The mediating role of need satisfaction. *Journal of Applied Sport Psychology*, 26, 182-196. doi: 10.1080/10413200.2013.819392
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). United States of America: Pearson Education, Inc.
- Tejeda, M. J., Scandura, T. A., & Pillai, R. (2001). The MLQ revisited: Psychometric properties and recommendations. *The Leadership Quarterly*, 12(1), 31-52.
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1-10. doi:10.1007/bf02291170
- Turnnidge, J., & Côté, J. (2019). Observing coaches' leadership behaviours. The development of the Coach Leadership Assessment System (CLAS). *Measurement in Physical Education and Exercise Science*, 1-13. doi:10.1080/1091367X.2019.1602835
- Zwingmann, I., Wegge, J., Rudolf, M., Richter, P., Wolf, S., & Schmidt, M. (2014). Is transformational leadership healthy for employees? A multilevel analysis in 16 nations. *Zeitschrift für Personalforschung*, 28(1-2), 24-51. doi: 10.1688/ZfP-01