

5-2011

# Situational Assessment on Leadership - Student Assessment (SALSA©): An Evaluation of the Convergent Validity with Multi-Source Feedback in Division I Intercollegiate Athletics

David Normansell

Western Kentucky University, david.normansell231@topper.wku.edu

Follow this and additional works at: <http://digitalcommons.wku.edu/theses>



Part of the [Industrial and Organizational Psychology Commons](#), and the [Personality and Social Contexts Commons](#)

---

## Recommended Citation

Normansell, David, "Situational Assessment on Leadership - Student Assessment (SALSA©): An Evaluation of the Convergent Validity with Multi-Source Feedback in Division I Intercollegiate Athletics" (2011). *Masters Theses & Specialist Projects*. Paper 1050.  
<http://digitalcommons.wku.edu/theses/1050>

This Thesis is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Masters Theses & Specialist Projects by an authorized administrator of TopSCHOLAR®. For more information, please contact [topscholar@wku.edu](mailto:topscholar@wku.edu).



SITUATIONAL ASSESSMENT OF LEADERSHIP – STUDENT ASSESSMENT  
(SALSA©): AN EVALUATION OF THE CONVERGENT VALIDITY WITH MULTI-  
SOURCE FEEDBACK IN DIVISION I INTERCOLLEGIATE ATHLETICS

A Thesis  
Presented to  
The Faculty of the Department of Psychology  
Western Kentucky University  
Bowling Green, Kentucky

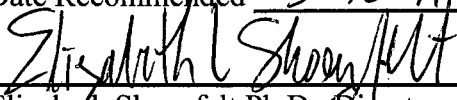
In Partial Fulfillment  
Of the Requirements for the Degree  
Master of Arts

By  
David Normansell

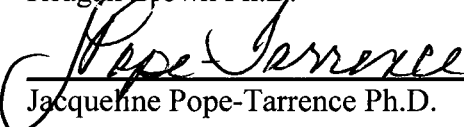
May 2011

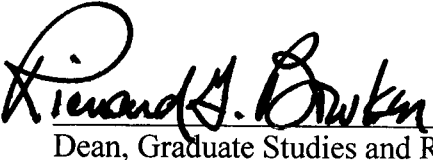
**SITUATIONAL ASSESSMENT OF LEADERSHIP – STUDENT ASSESSMENT  
(SALSA©): AN EVALUATION OF THE CONVERGENT VALIDITY WITH  
MULTI-SOURCE FEEDBACK IN DIVISION I INTERCOLLEGIATE  
ATHLETICS**

Date Recommended 5-10-11

  
Elizabeth Shoenfelt Ph.D., Director of Thesis

  
Reagan Brown Ph.D.

  
Jacqueline Pope-Tarrence Ph.D.

 May 16, 2011  
Dean, Graduate Studies and Research Date

## Contents

Abstract.....	vi
Introduction.....	1
Situational Judgment Tests .....	2
SJT Development.....	2
SJT Response Options .....	3
Threats to SJT Effectiveness.....	4
Benefits of SJT.....	4
Situational Assessment of Leadership – Student Assessment (SALSA®).....	6
Multi-Source Feedback.....	7
Psychometric Properties of MSF .....	8
Pros and Cons to MSF .....	8
Selection of Raters .....	9
Application and Impact of Feedback.....	10
Leadership in Sports .....	12
Youth Leadership Development Philosophy .....	13
The Current Research .....	15
Hypotheses.....	15
Method .....	18
Participants.....	18
Materials and Procedures.....	18
Results.....	19
Discussion.....	25

Additional Findings .....	26
Practical Application of Multi-Source Feedback.....	27
Limitations .....	28
Directions for Future Research .....	29
Conclusions.....	29
References.....	31
Appendix A.....	37
Appendix B.....	38
Appendix C.....	39
Appendix D.....	42

List of Tables

Table 1: Correlations Between Ratings from Different Sources on SALSA©  
Dimensions ..... 23

Table 2: Results for Self-Ratings Compared to Peer and Coach Ratings for 8 Dimensions  
of Leadership ..... 24

SITUATIONAL ASSESSMENT OF LEADERSHIP – STUDENT ASSESSMENT (SALSA©): AN EVALUATION OF THE CONVERGENT VALIDITY WITH MULTI-SOURCE FEEDBACK IN DIVISION I INTERCOLLEGIATE ATHLETICS

David Normansell

May 2011

49 Pages

Directed by: Drs. Elizabeth L. Shoenfelt, Reagan Brown, Jacqueline Pope-Tarrence

Department of Psychology

Western Kentucky University

The current study assessed the convergent validity of the Situational Assessment of Leadership – Student Assessment (SALSA©) and multi-source ratings in Division I intercollegiate athletic teams. Identified student-athlete team leaders were asked to complete the SALSA©, which assesses eight dimensions of leadership. By assessing the relationship between SALSA© scores and multi-source ratings (i.e., self, teammate, and coach), of the same eight leadership dimensions, a unique multi-dimensional perspective of leadership is revealed. Results indicated a significant positive relationship between overall SALSA© scores and overall self and coach performance ratings. Overall SALSA© scores also were significantly correlated with Overall Leadership Effectiveness peer-ratings. The dimension of Problem Solving/Innovation was significantly correlated with self-, peer-, and coach-ratings. Influencing Others and Communication SALSA© scores were significantly correlated with the self rating for their respective dimension. Self-ratings were significantly higher than any other source of rating. This study further validates the effectiveness of the SALSA© to identify and predict leadership behavior.



## Introduction

In 2009, Shoenfelt developed a situational judgment test called the Situational Assessment of Leadership – Student Assessment (SALSA©) that assessed student leadership skills on eight established leadership dimensions (i.e., Organizing/Visioning/Planning; Consideration/Team Skills; Problem Solving/Innovation; Influencing Others; Communication; Drive/Results Orientation; Tolerance for Stress; and Integrity/Ethics; Arthur, Day, McNelly, & Edens, 2003). The SALSA© scores obtained in initial studies were validated with leadership scores from the Assessment Center of the Center for Leadership Excellence at Western Kentucky University (Grant, 2009).

Leadership is important to any organization's success. A leader provides vision for the organization and the determination to turn that vision into a reality. A leader must be able to influence others to work together toward a common goal. Leaders must have integrity and a strong moral code to face and persevere through countless challenges. Understanding that challenges will be faced and mistakes will be made, a leader must be able to control his/her affective state and remain emotionally collected and focused on the task at hand.

The current study focused on the leadership skills of Division I intercollegiate athletes. Student-athletes completed the Situational Assessment of Leadership – Student Assessment (SALSA©) and rated their leadership ability on the eight leadership dimensions assessed by SALSA©. Coaches and teammates also were asked to rate the student-athlete's leadership skills on the same eight dimensions. SALSA© scores and these three sets of ratings (i.e., self, teammate, and coach) were correlated to assess convergent validity.

I will first review the literature on situational judgment tests. Next, I will address the development of the SALSA©. I will then discuss the benefits of multi-source feedback. The final section will describe the importance of leadership in a sports context.

### *Situational Judgment Tests*

Situational judgment tests (SJT) have been used for decades to quickly and effectively evaluate and/or predict future performance (McDaniel & Ngyuen, 2001). SJTs consist of a series of hypothetical job-related (typically) scenarios and statements of behavioral responses to each scenario to predict how the employee (in most cases) will respond to similar situations on the job. SJTs are intended to measure typical or optimal performance (Lievens, Peeters & Schollaert, 2008). An SJT measures typical performance by asking test takers what they “would do” in a given situation. Optimal performance is assessed by asking test takers what they “should do” in the situation (McDaniel & Ngyuen, 2001).

### *SJT Development*

The process by which SJTs are constructed begins with identifying item content. Weekley, Ployhart, and Holtz (2006) stated that, in most cases, subject matter experts (SMEs; e.g., job incumbents, supervisors, or leading professionals in the field) identify critical incidents that occur on the job. Identifying critical incidents usually follows the three-part or A-B-C format (i.e., Antecedent – Behavior – Consequence). The antecedent refers to the circumstances that led to the critical incident. The behavior refers to the action that was taken and whether it was effective or ineffective; the consequence describes the results of the behavior.

### *SJT Response Options*

Methods for determining response options to critical incidents vary considerably depending on the context. One must consider the source of the response options, complexity, and fidelity of the options, and the degree to which respondents could fake correct answers (Weekley et al., 2006).

When determining the source of the response options, most developers use SMEs to formulate a list of possible behaviors (Motowidlo, Hanson, & Crafts, 1997). Other developers rely on themselves to devise response options (Weekley & Ployhart, 2005). No research, however, has determined which strategy is most effective (Weekley et al., 2006). Using both the organizational and tacit knowledge of the SMEs to construct response options might be more beneficial in terms of realism of responses. A large number of SMEs responding results in a broad range of responses, which is beneficial when choosing the final response options.

Complexity is important to note when considering both the item and response level (Weekley et al., 2006). As with any instrument, it is important to consider respondent reading level when formulating the items and the response options. If the items and response options contain too much organization-specific information, it will be difficult to generalize results or to use the instrument outside of the organization. One should, however, be aware that organization-specific information is helpful in establishing psychological fidelity. In responding to critical incidents (i.e., specific situations that occur on the job), respondents must consider the psychological/cognitive demands of the job to determine the correct course of action. It is no surprise that research has shown that employment status affects SJT performance (Weekley, Ployhart,

& Harold, 2004). Incumbents have been shown to score higher than applicants on cognitive ability SJTs. This is presumably due to tacit knowledge gained on the job (Weekley et al., 2004). For the current study, differences on SALSA© scores based on duration of time on the team would have been assessed; however, due to the coaches' choice on who should be assessed, only three sophomores and no freshmen were chosen. As such, no calculations were made assessing this possible difference.

#### *Threat to SJT Effectiveness*

Psychological fidelity enhances an SJT's effectiveness, however, other factors can hinder its usefulness. The ability for respondents to fake responses is detrimental to the sensitivity of the instrument. Faking is the ability to choose the correct answer based on the transparency of the question and response options (i.e., fakers choose the correct response without knowledge of the correct response). These test takers are said to have "faked" their responses because they responded in a way that, even though correct, is different from the way they would have responded to that situation on the job. Instruments are designed to discriminate between high and low performers. If low performers can answer correctly the same questions as high performers, the test lacks sensitivity.

#### *Benefits of SJT*

When used correctly, SJTs are very versatile in their ability to assess a wide range of constructs, including leadership. An example of a leadership SJT is the Leadership Skills Assessment (Bergman, Drasgow, Donovan, Henning & Juraska, 2006). SJT scores have been shown to correlate with predictors of job performance, cognitive ability, and three of the Big Five Factors of personality. McDaniel, Hartman, Whetzel, and Grubb

(2007) found that cognitive ability was strongly correlated with SJTs that assess optimal performance (i.e., asking what one “should do”). Typical performance based SJTs (i.e., asking what one “would do”) strongly correlate with personality constructs. SJTs have been shown to significantly correlate with predictors of emotional stability ( $r = .31$ ), agreeableness ( $r = .25$ ), and conscientiousness ( $r = .26$ ) (McDaniel & Nguyen, 2001). McDaniel, Morgeson, Finnegan, Campion, and Braverman (2001) found SJTs to significantly correlate with predictors of job performance ( $p = .34$ ) and cognitive ability ( $p = .46$ ).

In a 2010 meta-analysis, Christian, Edwards, and Bradley discussed the importance of properly identifying the constructs being measured by SJTs. Too often researchers focus solely on an SJTs composite score and its relation to job performance, while rarely focusing on the individual constructs that that test measures. Arthur and Villado (2008) stated that the goal of using SJTs should not only be to predict future job performance, but also to assess why the test, more specifically, the construct, predicted performance. Christian et al. indicated that it was important to understand why a construct predicted performance because it enables researchers to compare different SJTs based on the constructs they measured. To appropriately assess the psychometric properties (e.g., reliability, validity, and differences between subgroups) on SJTs, one must first identify the construct or constructs that a particular test measures. Results of the meta-analysis (Christian et al.) indicated that most SJTs measured either leadership or interpersonal skills. Leadership and team work skills yielded the strongest validity for job performance. These results indicated just how important leadership skills are to an organization’s success.

Situational judgment tests have been used for decades and have assessed numerous constructs, including leadership. Due to their versatility and sound psychometric properties, SJTs will continue to be widely used and remain a popular assessment technique. The SJT used in the current study will now be further discussed.

*Situational Assessment of Leadership – Student Assessment (SALSA©)*

The Situational Assessment of Leadership – Student Assessment (SALSA©) is a situational judgment test that measures eight dimensions of leadership. The leadership dimensions used by the SALSA© consist of seven dimensions that Arthur et al. (2003) identified as the most common dimensions of leadership. These dimensions are Organizing/Visioning/Planning, Consideration/Team Skills, Problem Solving/Innovation, Influencing Others, Communication, Drive/Results Orientation, and Tolerance for Stress. An eighth dimension, Integrity/Ethics, was added (Shoenfelt, 2009). Descriptions of these dimensions can be found in Appendix A.

Similar to other cognitive ability SJTs, SALSA© presents a challenging situation or conflict to the test taker then asks him or her to resolve the incident by choosing one of four possible responses. There are 120 items divided among the eight dimensions.

Grant (2009) found that the SALSA© has an internal consistency of  $\alpha = .91$ . The leadership dimensions of the SALSA© were matched with dimensions previously used for the Western Kentucky University Center for Leadership Excellence Assessment Center. Dimension scores from the SALSA© were correlated with assessment center scores; convergent validity was statistically significant but low in magnitude ranging from  $r = .28$  to  $r = .44$ . Weekley et al. (2006) stated that, although SJTs are frequently broken down by dimensions, few dimensional subscores are ever reported due to

multidimensionality and poor psychometric properties. The SALSA©, however, is unique in the fact that its dimensional subscores are not highly correlated with each other; the average bivariate correlation between dimensions is  $r = .52$  (Slack, 2010). One possible explanation is that other SJTs tend to assess relatively narrow constructs (e.g., cognitive ability), which in turn results in dimensional subscores (e.g., problem-solving skills, metacognitive skills) being highly correlated with each other. The SALSA©, on the other hand, is designed to assess the construct of leadership, which is relatively broad. The dimensions assessed in SALSA© (e.g., Problem Solving/Innovation, Influencing Others, and Drive/Results Orientation) have proven to be diverse enough to allow dimension subscores to be assessed and reported during the evaluation of overarching leadership ability.

That the eight dimensions are not highly correlated with each other suggests each dimension explains unique variance in the construct of leadership. Each dimension is not only essential to defining leadership, but also in identifying and developing leadership ability. In the current study, these leadership dimensions were assessed multiple times for each student-athlete leader. By obtaining rating in each dimension from multiple sources (i.e., self, teammate, and coach) for each student-athlete leader, a more comprehensive evaluation of leadership was conducted. The following section further discusses the uses of multi-source feedback.

#### *Multi-Source Feedback*

The use of multi-source feedback (MSF) offers a unique opportunity to provide a comprehensive multi-level perspective of an individual's performance. The use of multiple raters (e.g., superiors, peers, self, and subordinates) provides the target

individual with insight into how others perceive their performance (Foster & Law, 2006). MSF requires two or more raters from different levels in an organization (e.g., supervisor and peer ratings; Dalessio, 1998). Although frequently used synonymously in the literature, 360-degree feedback and MSF are not synonymous (Foster & Law, 2006). Three-hundred-sixty degree feedback includes all four possible levels of raters – supervisors, peers, self, and subordinates/clients/customers. Three-hundred-sixty degree feedback is a form of MSF. For the current study, as there were no subordinate ratings, the term MSF was used.

#### *Psychometric Properties of MSF*

The psychometric properties of multi-source feedback are one of the key reasons for its popularity. Internal consistency is normally sufficient with coefficient alphas in the .70 range for each dimension of assessment (Van Velsor & Leslie, 1991). Conway and Huffcutt (1997) investigated within-source agreement (e.g., among subordinates, among peers, and among supervisors) and found only moderate relationships: subordinates (.27), peer (.39), and supervisor (.51). When investigating inter-rater agreement, the discrepancies between sources may be related to each groups' unique perspective of the participant's behavior (Dalessio, 1998). Low inter-rater agreement does not necessarily suggest poor psychometric properties; rather, it likely reflects genuine differences in the way different groups perceive an individual's behavior.

#### *Pros and Cons to MSF*

There are a variety of benefits to using multi-source feedback. One benefit is that MSF provides a broader view of overall performance. Second, MSF allows the one being evaluated to recognize how his/her actions are perceived by others. For example, a study



conducted assessing managerial performance using single source feedback and MSF found that managers were more accepting of results that utilized multiple sources rather than a single source (Garavan, Morley, & Flynn, 1997). Last, individual ratings are normally kept confidential, thus providing evaluators the chance to provide honest feedback with regard to the individual's performance.

There are several potential areas of concern regarding multi-source feedback. Inflated self-ratings are very common. Researchers attempt to identify the effect of inflated self-ratings by assessing rater agreement. Self ratings tend to be higher than ratings from any other source (i.e., Bass & Yammarino, 1991; Brutus, Fleenor, & McCauley, 1997; Harris & Schaubroeck, 1988). Participants tend to exaggerate their abilities, especially during evaluations. Normally used as a defensive technique in response to the fear of being perceived as less than adequate, participants tend to inflate self-ratings. It is not uncommon for participants, especially if they feel they are being evaluated, to slightly elevate self-ratings in an attempt to improve their image/worth in the face of others.

#### *Selection of Raters*

How raters are selected is an important issue that has resulted in considerable controversy. When using MSF, the method in which raters are chosen presumably can lead to biased results. If raters are chosen at random, there is a chance that little or no interaction has occurred between the rater and targeted individual (i.e., the one being assessed). The ratings would be meaningless given the lack of opportunity to observe the individual. Scores would not reflect actual ability or performance. On the other hand, if participants are able to choose their own raters, ratings are susceptible to significant

positive bias (MaCarthy & Garavan, 2001). Having no perfect solution, most MSF evaluations allow participants to list perspective raters that, presumably, are in the best position to make accurate assessments (Atwater, Brett, & Charles, 2007). Research indicates that if participants chose their own raters, participants consider them more credible and feedback is perceived more favorably (Smither, London, & Reilly, 2005). Becton and Schraeder (2004) added that the opportunity to select raters will increase acceptance of the feedback. For the current study, bias would have been minimal given the nature of the situation; that is, that all participants were members of the same intercollegiate team and would have had significant interaction with each other and opportunity to observe the leadership ability of the other players.

Research on the number of raters needed for an accurate assessment has had mixed results. Some researchers suggest a minimum of three raters per level (Garavan et al., 1997), while others suggest a minimum of six raters (Wood, Hassell, Whitehouse, Bullock, & Wall, 2006) to ensure accurate results. In the current study, the average number of peer ratings was fourteen. All three teams had a minimum of three coaches provide the coach ratings.

#### *Application and Impact of Feedback*

Following the formal evaluation (i.e., multi-source data collection and data analyses), feedback is given to the participant based on the rating results. By providing a multi-source view of one's behavior, those being evaluated are made aware of discrepancies between their own perceptions and the perceptions of others (e.g., supervisor, peer, and subordinate). Research has indicated that managers often are unaware that these discrepancies exist (Metcalf, 1998). When made aware, however,

managers are more likely to seek out ways to develop deficit skills as a way of reducing the discrepancies (Atwater et al., 2007). This increased motivation provides the driving force behind organization change and improvement. In a recent study, seventy percent of surveyed companies that use multi-source feedback evaluations utilized the results for developmental improvements rather than for administrative decision making (Atwater et al., 2007).

Different personality characteristics have been shown to impact the effectiveness of multi-source feedback. Emotional stability, extroversion, conscientiousness, and openness to experience can all affect reactions to MSF (Smither, London, & Reilly, 2005). Smither, London, and Richmond (2005) found that managers who are more emotionally stable have increased motivation to use the feedback to improve performance following MSF. Also, Atwater and Brett (2005) found that participants with low emotional stability reported more negative emotions regardless of the feedback direction (i.e., positive or negative). Smither, London, & Richmond (2005) found that participants with high levels of extroversion were more likely to ask for additional feedback regarding their performance. Following MSF, conscientious participants felt obligated to use feedback and were more likely to participate in activities to develop perceived deficiencies (Smither, London, & Richmond, 2005). Furthermore Dominick, Reilly, and Byrne (2004) found that following MSF, conscientiousness and openness to experience were positively related to improved performance.

Multi-source feedback is a very popular technique that provides participants with the opportunity to see how others perceive their actions. Based on its sound psychometric properties, its versatility with regard to rater selection, and its positive relationship with

several personality dimensions, MSF can be a valuable asset in effectively evaluating the construct of leadership.

### *Leadership in Sports*

Leadership in sports is displayed when an athlete takes on a formal (captain) or informal leadership role and influences his or her teammates to work together for a common goal (Loughead, Hardy, & Eys, 2006). Glenn and Horn (1993) suggested that the presence of a leader is a vital component of any successful sports team. Identifying leadership skills in athletes provides coaches with insight on how to develop players with varying levels of leadership skills.

Parents, coaches, and school administrators often associate participation in sports as playing a critical role in the development of essential life skills, particularly leadership (Gould, Carson, Fifer, Lauer, & Benham, 2009). Voelker, Gould, and Crawford (2010), however, warn that this conception may be ill advised based on their research which examined the leadership roles of 13 high school student-athlete captains. Results indicated that none of the 13 captains received any additional leadership training from their coaches. Furthermore, the research suggested that the captains did not receive proper mentorship from their coaches and were not given the opportunity to express leadership skills such as leading team meetings (Voelker et al.). This suggests that little if any additional leadership skills were learned as the result of being team captain.

Still, this does not discredit the fact that countless studies have suggested the need to develop leadership abilities in student-athletes (e.g., Gould, Smith, White, & Chung, 2006). Leadership ability has been shown numerous times to positively impact athletic performance. Callow, Smith, Hardy, Arthur, and Hardy (2009) found some leadership

behaviors (fostering acceptance of group goals and promoting teamwork, high performance expectations, and individual consideration) significantly predicted task cohesion. Fostering acceptance of group goals and promoting teamwork was also shown to predict social cohesion. Van Vianen and De Dreu (2001) found a significant relationship between task and social cohesion and performance.

### *Youth Leadership Development Philosophy*

In a recent article, the Institute for the Study of Youth Sports (ISYS) outlined its philosophy on youth leadership development (Gould & Voelker, 2010). The philosophy is based on leadership research with adults (Anderson, 2007) as well as children (Martinek & Hellison, 2009; Van Linden & Fertman, 1998) to ensure a comprehensive understanding of the importance of the development of leadership skills (Gould & Voelker, 2010). While the philosophy is tailored to high school athletics, its principles should generalize to intercollegiate athletics as well.

Gould and Voelker (2010) outlined ISYS's leadership philosophy as seven principles. The first ISYS leadership principle stems from the basic understanding that leadership is the process of working with a group to identify and achieve personal and/or group goals. The second principle takes into account additional factors (e.g., leader characteristics, follower characteristics, and situational influences) that influence the effectiveness of the leader. The third principle recognizes that individual characteristics and previous experiences will facilitate the acquisition of leadership skills in some individuals more so than others. The fourth principle identifies the stages of leadership development starting with awareness and basic skill development (e.g., making eye contact and using proper speech) and ending with the understanding and manipulation of

one's environment (e.g., knowledge of interpersonal dynamics, eliciting behaviors from followers). Recognizing that learning to lead stems from multiple areas is the fifth principle mentioned. Ways in which leadership skills may be learned include previous experience, trial and error, mentorship, and formal education. The sixth principle states that leadership can only be developed if the athlete is given the opportunity to perform the learned skills. This is a crucial component considering that coaches have been shown to often minimize the leadership responsibilities of team captains (Voelker et al., 2010). The seventh and final principle ties into Voelker and his colleagues' findings by stating that leadership ability and responsibility are underutilized in sports. Captains are not being given the opportunity to take on a real leadership role on the team. Captainship has become a figurehead position holding no legitimate leadership power. Still, Gould and Voelker stressed that captainship, if utilized appropriately, can provide coaches with the opportunity to teach athletes essential leadership skills.

Coaches must recognize that leadership skills will not be developed through athletic participation alone. These skills need to be taught and exercised to establish lasting effects. Often coaches become frustrated with the lack of leadership on their teams; however, few proactively seek to develop these skills in their athletes (Gould & Voelker, 2010).

Advancing the previous research regarding the need to develop leadership abilities in student-athletes (e.g., Gould et al., 2006), the current study provided coaches with the opportunity to proactively identify leadership ability in their student-athletes using multiple sources (i.e., self, teammate, and coaches). MSF has been demonstrated to have utility for individual development as well as to motivate individual development

(Atwater et al., 2007). This broad understanding of how the team perceives an individual's leadership ability is the first step for a coach to recognize and develop that player's leadership skills. Coaches will have a better understanding of how to develop an athlete's areas of weakness and utilize areas of strength to increase the success of the team.

Following ISYS's philosophy on youth leadership development (Gould & Voelker, 2010), the current study hoped to expand the understanding of leadership development by taking into account multiple sources of perceived leadership ability. Although research suggests that team leaders have not received proper leadership guidance from their coaches in the past (Voelker et al., 2010), the current research sought to evaluate the leadership abilities in student-athletes with the hope that coaches will use this information to further the development of their athletes' leadership abilities.

#### *The Current Research*

With the cooperation of Western Kentucky University intercollegiate athletic teams, the current research assessed the convergent validity of multi-source ratings with SALSA© scores. Student-athletes, who were identified as leaders by their coaches, were assessed by SALSA© on eight dimensions of leadership. Teammates and coaches rated the student-athlete on the same dimensions of leadership; self-ratings were also completed. These scores and ratings were then correlated.

#### *Hypotheses*

It was expected that SALSA© scores and multi-source ratings would demonstrate convergent validity both for composite scores and for dimension scores. SALSA© composite scores were correlated with composite performance ratings and Overall

Leadership Effectiveness ratings from each source (i.e., self, teammate, and coach's ratings).

Hypothesis 1a: SALSA© composite scores will positively correlate with composite performance ratings from each source (i.e., self, teammate, and coach).

Hypothesis 1b: SALSA© composite scores will positively correlate with Overall Leadership Effectiveness ratings from each source (i.e., self, teammate, and coach).

The eight dimensions of SALSA© (Organizing/Visioning/Planning; Consideration/Team Skills; Problem Solving/Innovation; Influencing Others; Communication; Drive/Results Orientation; Tolerance for Stress; and Integrity/Ethics) parallel the eight dimensions on the performance rating form. Each SALSA© dimension was correlated with its corresponding dimension ratings from each source (i.e., self, teammate, and coach).

Hypotheses 2a-h: Each dimension score on SALSA© will positively correlate with the corresponding performance ratings from each source (i.e., self, teammate, and coach).

Hypothesis 2a: SALSA© dimension

Organizing/Visioning/Planning will positively correlate with performance rating dimension Organizing/Visioning/Planning.

Hypothesis 2b: SALSA© dimension Consideration/Team Skills will positively correlate with performance rating dimension Consideration/Team Skills.



Hypothesis 2c: SALSAC dimension Problem Solving/Innovation will positively correlate with performance rating dimension Problem Solving/Innovation.

Hypothesis 2d: SALSAC dimension Influencing Others will positively correlate with performance rating dimension Influencing Others.

Hypothesis 2e: SALSAC dimension Communication will positively correlate with performance rating dimension Communication.

Hypothesis 2f: SALSAC dimension Drive/Result Orientation will positively correlate with performance rating dimension Drive/Result Orientation.

Hypothesis2g: SALSAC dimension Tolerance for Stress will positively correlate with performance rating dimension Tolerance for Stress.

Hypothesis2h: SALSAC dimension Integrity/Ethics will positively correlate with performance rating dimension Integrity/Ethics.

Previous research has indicated that self-ratings tend to be higher than ratings from other sources (i.e., Bass & Yammarino, 1991; Brutus et al., 1997; Harris & Schaubroeck, 1988). Self-ratings were compared to ratings from teammates and coaches to determine if self-ratings in the current study were inflated.

Hypotheses 3: Self-ratings will be higher than ratings from other sources on all eight SALSA© dimensions.

Cognitive-ability SJTs tend to use organization-specific information in an attempt to increase the instrument's psychological fidelity. Using this organization-specific information to one's advantage, job incumbents have been shown to do better on SJTs than job applicants (Weekley et al., 2004). This higher performance is presumably due to higher cognitive ability resulting from tacit knowledge gained on the job. While the SALSA© is not sport specific, it was hypothesized that the more time spent on the team, the greater one's learned and perceived leadership ability.

Hypothesis 4a: Seniors and juniors will score higher than freshmen and sophomores on SALSA©.

Hypothesis 4b: Seniors and juniors will be rated higher than freshmen and sophomores by coaches and teammates.

## **Method**

### *Participants*

The Western Kentucky University (WKU) women's volleyball, soccer, and softball teams were participants in my study (N = 18). The average age of participants was 20.53 years ( $SD = .87$ ); all student-athlete leader participants were female.

### *Materials and Procedure*

The study was conducted during Christmas break, January term, and the first part of spring semester to ensure that all members of the teams had equal and ample opportunity to participate. Coaches who agreed to participate were asked to identify the student-athletes they wanted to participate in the leadership evaluation. For each student-

athlete, the coach completed a form that asked for the student-athlete's name and names of coaches and teammates that would serve as raters. Participating student-athletes were asked to complete the SALSA© online via Blackboard. SALSA© took about one hour to complete. Items on the SALSA© ask participants to identify the most appropriate action a leader should take in particular leadership situations. Each student-athlete leader was asked to complete self-ratings assessing the same eight dimensions as the SALSA© as well as an Overall Leadership Effectiveness rating. Identified teammates and coaches were then asked to rate the student-athlete on the same eight dimensions of leadership as well as on the Overall Leadership Effectiveness rating. For each leadership dimension, the rating form provided raters with a definition and behavioral examples of the dimension and a 5-point rating scale. A sample of the performance dimension rating form may be found in Appendix B.

Following the data analyses, individualized feedback packets enabled student-athletes to see their overall SALSA© score and an Overall Leadership Effectiveness rating, as well as SALSA© scores and ratings on each of the 8 dimensions of leadership. Narrative information for each SALSA© dimension also was provided. Student-athletes were provided information about where they fell relative to other students on each dimension. Coaches were provided with the student-athlete information for developmental purposes.

## **Results**

An overall SALSA© score was calculated for each individual by summing across the eight dimension scores. Two overall ratings were calculated for each student-athlete from the peer and coach ratings. The first overall rating from the peers was calculated by

averaging ratings from each peer across dimensions. The second overall rating from the peers was calculated by averaging the peer Overall Leadership Excellence (OLE) ratings. For the coaches, the first overall rating was calculated by averaging ratings from each coach across dimensions. The second overall rating from the coaches was calculated by averaging the coaches Overall Leadership Excellence ratings. For peer ratings, the averaged ratings across dimensions was  $M = 3.49$  ( $SD = .40$ ) and the average Overall Leadership Excellence ratings was  $M = 3.44$  ( $SD = .52$ ). For coach-ratings, the averaged ratings across dimensions was  $M = 3.69$  ( $SD = .40$ ) and the average Overall Leadership Excellence ratings was  $M = 3.50$  ( $SD = .55$ ).

Hypothesis 1 predicted that SALSA© scores and multi-source ratings would demonstrate convergent validity both for composite scores and for dimension scores. SALSA© composite scores were correlated with composite performance ratings and Overall Leadership Effectiveness ratings from each source (i.e., self, teammate, and coach ratings). Hypothesis 1a stated that the total SALSA© score would positively correlate with self, peer, and coach total dimension ratings. This hypothesis was partially supported (self ( $r = .53, p = .01$ ), peer ( $r = .36, p = .07$ ), and coach ratings ( $r = .51, p = .01$ )). Hypothesis 1b stated that the total SALSA© score would positively correlate with self, peer, and coach OLE dimension ratings. This hypothesis was supported (self ( $r = .56, p < .01$ ), peer ( $r = .42, p = .04$ ), and coach ratings ( $r = .52, p = .01$ )).

In Hypotheses 2a-h, it was predicted that each dimension score on SALSA© would positively correlate with the corresponding performance rating. This was tested by correlating each SALSA© dimensions score with the corresponding dimension rating from each rating source. Results indicated a significant correlation between the Problem

Solving/Innovation SALSA© score and the self, peer, and coach ratings for the Problem Solving/Innovation dimension (self ( $r = .50, p = .01$ ), peer ( $r = .64, p < .01$ ), and coach ( $r = .65, p < .01$ ), thus supporting Hypothesis 2c. There was also a significant correlation between the Influencing Others SALSA© score and the self rating for that dimension ( $r = .64, p < .01$ ), partially supporting Hypothesis 2d. The only other significant correlation was between the Communication SALSA© score and the Communication self rating ( $r = .41, p < .05$ ), partially supporting Hypothesis 2e. It is interesting to note, of the nine ratings evaluated (i.e., eight SALSA© dimensions and one OLE dimension), eight dimensions had a significant correlation between coach and peer ratings; only Tolerance for Stress did not. For the dimension of Consideration/Team Skills, self ratings significantly correlated with coach ratings ( $r = .43, p < .05$ ). For the dimension of Communication, self ratings significantly correlated with peer ratings ( $r = .54, p < .05$ ). See Table 1 for the specific correlations between raters. Refer to Appendix D for all correlations between ratings.

Hypothesis 3a-h predicted that self-ratings on each of the eight leadership dimensions would be significantly higher than ratings in that dimension by peers and coaches. Overall self-ratings were significantly higher ( $M = 3.94, SD = .48$ ) than overall peer ratings ( $M = 3.48, SD = .40; t(17) = 3.65, p < .01$ ) and overall coach ratings ( $M = 3.69, SD = .40; t(17) = 2.03, p < .05$ ). For specific relationships between raters by dimensions, see Table 2.

Table 1

*Correlations Between Ratings from Different Sources on SALSA© Dimensions*

	Self	Peer	Coach
Organizing/Planning/Visioning			
Self	--	0.29	0.24
Peer		--	0.72**
Coach			--
Consideration/Team Skills			
Self	--	0.29	0.43*
Peer		--	0.73**
Coach			--
Problem Solving/Innovation			
Self	--	0.37	0.37
Peer		--	0.51*
Coach			--
Influencing Others			
Self	--	0.03	0.33
Peer		--	0.47*
Coach			--
Communication			
Self	--	0.54*	0.35
Peer		--	0.71**
Coach			--
Drive/Results Orientation			
Self	--	0.38	0.40
Peer		--	0.68**
Coach			--
Tolerance for Stress			
Self	--	0.12	0.31
Peer		--	0.39
Coach			--
Integrity/Ethics			
Self	--	0.25	0.06
Peer		--	0.83**
Coach			--
Overall Leadership Excellence			
Self	--	0.31	0.22
Peer		--	0.60**
Coach			--

\* $p < .05$ , \*\* $p < .01$

Table 2

*Results for Self-Ratings Compared to Peer and Coach Ratings for 8 Dimensions of Leadership*

Dimension	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>
Organizing/Planning/Visioning				
Self	4.06	0.80	-	-
Peer	3.45	0.61	17	2.99**
Coach	3.68	0.46	17	1.95*
Consideration/Team Skills				
Self	4.06	0.64	-	-
Peer	3.58	0.54	17	2.87**
Coach	3.67	0.73	17	2.23*
Problem Solving/Innovation				
Self	3.89	0.76	-	-
Peer	3.37	0.58	17	2.86**
Coach	3.56	0.45	17	1.86*
Influencing Others				
Self	3.61	0.85	-	-
Peer	3.25	0.45	17	1.60
Coach	3.59	0.57	17	.09
Communication				
Self	4.00	0.91	-	-
Peer	3.39	0.63	17	3.35**
Coach	3.42	0.59	17	2.74**
Drive/Results Orientation				
Self	4.22	0.55	-	-
Peer	3.81	0.52	17	2.90**
Coach	3.92	0.65	17	1.93*
Tolerance for Stress				
Self	3.61	0.85	-	-
Peer	3.31	0.51	17	1.35
Coach	3.31	0.61	17	1.43
Integrity/Ethics				
Self	4.11	0.58	-	-
Peer	3.71	0.59	17	2.35*
Coach	4.34	0.55	17	-1.26
Overall Leadership Excellence				
Self	3.94	0.42	-	-
Peer	3.44	0.52	17	3.85**
Coach	3.50	0.55	17	3.03**

Note: t-test compares to self-rating.

\* $p < .05$ , \*\* $p < .01$

Hypothesis 4 predicted that seniors and juniors would score higher on SALSA© and be rated higher by coaches and peers than freshmen and sophomores. Only two sophomores and no freshmen participated in the study; thus, those analyses were not conducted.

### **Discussion**

This study was intended to assess the convergent validity between SALSA© scores and multi-source feedback. The assessment compared SALSA© dimension scores with self, peer, and coach ratings. Results indicated that overall SALSA© scores significantly correlated with overall self and coach ratings. While peer overall ratings were not significantly correlated with overall SALSA© scores, peer OLE scores were significantly correlated with overall SALSA© scores. These results suggest that SALSA© is a good predictor of leadership for intercollegiate female athletes. That is, higher SALSA© scores are related to higher perceptions of leadership by one's self, peers, and coaches.

On the dimension level, Problem Solving/Innovation ratings were significantly correlated across all sources of ratings. That is, all sources of raters shared similar perceptions of the student-athlete's level of Problem Solving/Innovation skills. SALSA© scores for Influencing Others and Communication were significantly correlated with self-ratings but not with coach or peer-ratings. This finding raises some interesting questions. These results indicate that student-athletes were accurate in their perceptions of their communication and influencing skills as measured by SALSA©. These perceptions, however, were not reflected by peer and coach-ratings, the sources of which are ones with whom the student-athletes were presumably communicating and influencing. It is



important to note, however, that for the dimension of Influencing Others, there was no significant difference between sources of ratings. This finding indicates that all sources have the same perception of a student-athlete's level of influence on others. An explanation for these contradicting findings is not clear. For the dimension of Communication, there was a significant difference between both peer and coach ratings and self-ratings. Results indicate that although student-athletes possess an accurate view of their communication skills, these skills may not be exhibited to or perceived by their peers and coaches.

#### *Additional Findings*

For eight of the nine dimensions tested (eight SALSA© dimensions and OLE rating) there were significant correlations between coach and peer ratings. Only Tolerance for Stress proved to have no significant correlations between rating sources. Research supports that different ratings from different rating sources may be due to differences in opportunities to witness leadership dimensions. Perhaps in a sports team context, leaders are observed in various situations by all raters; that is, players and coaches share opportunities to observe leaders. For the dimension of Tolerance for Stress, ratings from all sources were among the lowest among the eight dimensions (i.e., self-rating was tied for lowest, peer-rating was second lowest, and coach-rating was lowest). It is important to note, that Tolerance for Stress was an area identified by all teams a priori as an area of concern and a reason for initially participating in the study.

As predicted, overall self-ratings were significantly higher than peer and coach ratings. On the eight dimension ratings, self-ratings were significantly higher than peer and coach-ratings with the exception of Influencing Others and Tolerance for Stress,

where there were no differences between rating sources. It is interesting to note that Integrity/Ethics self-ratings were not significantly different than coach-ratings. Although Integrity/Ethics was the second highest peer-rating, self-ratings were still significantly higher than peer ratings. The Integrity/Ethics rating was the highest coach-rating and coach-ratings were not significantly different than self-ratings.

### *Practical Application of Multi-Source Feedback*

Following the research study, each team sought out different ways to continue their leadership development. The women's soccer team was interested in differences found between sources of ratings; in particular, why self-ratings were higher than peer and coach-ratings. All soccer coaches emphasized the importance of honest and accurate self-ratings and suggested these discrepancies indicate that athletes were not holding themselves accountable for weaknesses in leadership ability. The coaches made it a point to emphasize that players must first be honest with themselves if they seek to improve.

As a younger team, the softball team's primary focus was team cohesion. The leaders on their team, identified a priori by the coach as the juniors and seniors, wanted to work on inspiring their younger teammates to work hard as they grow into a cohesive unit. Establishing trust takes time and will presumably strengthen as their spring season continues.

The volleyball coach was particularly interested in leadership identification and development. The leaders used the individualized feedback reports as a springboard to a "position partner" program. In the program, each player was assigned a teammate/partner to hold accountable for demonstrating mental toughness during the spring season (e.g., not getting frustrated, taking emotions out of the team, or assigning blame, etc.). Along

with the program, the coach conducted his own re-evaluation of leadership to serve as a follow-up to my research.

### *Limitations*

Small sample size was the biggest limitation in this study. The Type I error rate is assured for the significant results; however, Type II error may have occurred. That is, results that may reach significance with a larger sample size were not detected with the current small sample of 18.

Although participants were asked to be honest and accurate in their ratings before the rating forms were distributed, ratings errors may have occurred. There was no rater training, although rater training exercises may increase accuracy of ratings. Contrast effects refer to raters making ratings by comparing the targeted student-athlete with other individuals rather than to the standard. It is unknown whether contrast effects introduced error in the ratings. Halo error refers to when raters generalize from one dimension of leadership performance to other dimensions of leadership. Leniency refers to raters assigning ratings to a student-athlete that are higher than deserved, presumably based on friendships. Halo error and leniency were concerns because of the friendships that undoubtedly have formed between team members. However, peer ratings were significantly lower than self-ratings on six of the eight leadership dimensions and on the OLE rating, suggesting that leniency error was not occurring.

Another limitation to the current study is the use of a single-item measure to assess each leadership dimension. Research indicates that single-item measures should only be used if the single item reflects a homogeneous concept, as indicated by a high internal consistency coefficient ( $\alpha > .85$ ; Loo, 2002). The coefficient alpha for

SALSA© is  $\alpha = .91$  (Grant, 2009). Furthermore, in a parallel study to the current research, Tucker (2011) examined the relationship between ratings of the behavioral items that serve to define each dimension and the global rating for that dimension. Tucker correlated overall behavioral ratings (from each source) with global performance ratings (single-item ratings; from each source) on the same eight dimensions measured by SALSA©. His results indicated that for each source and across all dimensions there were significant relationships between the global rating and the behavioral ratings. Therefore, Tucker's results indicate that a single-item measure provided ratings very similar to the global ratings on the eight leadership dimensions measured by SALSA©. See Appendix C for the correlation table between overall ratings on each dimension for each source from Tucker.

#### *Directions for Future Research*

Future researchers should strive to increase the sample size and demographic variability of the sample. The participants in this study were all female with little demographic variability. The use of multi-source ratings in male sports teams as well as individual sports may provide interesting perspectives on leadership. Rater training exercises may increase accuracy of ratings by all sources.

#### *Conclusions*

This study assessed the convergent validity between SALSA© and multi-source ratings in Division I intercollegiate female athletic teams. While each team appeared receptive to the feedback, only time will tell if the student-athletes will use their new knowledge of leadership to improve in their respective leadership roles. After all the individualized feedback sessions were delivered, all teams expressed interest in further

leadership development. As a result, Dr. Shoenfelt offered a three-week mental skills development course to all teams. The course emphasized skills and techniques athletes could use to develop mental toughness. By knowing which areas of leadership need improvement (e.g., Tolerance for Stress), student-athletes were perhaps more inclined to practice and use these newly-acquired techniques to improve their leadership abilities.

This study compared multiple ratings to a situational judgment test designed to assess leadership ability on eight dimensions. This provided a unique multi-method perspective of leadership. Findings suggest a positive correlation between overall SALSA© scores and multi-source ratings. This study further validates the Situational Assessment of Leadership-Student Assessment's ability to accurately identify and predict leadership ability.

## References

- Anderson, N. B. (Ed.). (2007). Leadership [Special issue]. *American Psychologist*, 62, 1-47.
- Arthur, Jr., W., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology*, 53, 125-154.
- Arthur W., Jr., & Villado, A. J. (2008). The importance of distinguishing between constructs and methods when comparing predictors in personnel selection and research in practice. *Journal of Applied Psychology*, 93, 435-442.
- Atwater, L. E. & Brett, J. F. (2005). Antecedents and consequences of reactions to developmental 360-degree feedback. *Journal of Vocational Behavior*, 66, 532-548.
- Atwater, L. Brett, J., & Charles, A. (2007). Multi-source feedback: Lessons learned and implications for practice. *Human Resource Management*, 46, 285-307.
- Bass, B. M., & Yammarino, F. J. (1991). Congruence of self and others' leadership ratings of naval officers for understanding successful performance. *Applied Psychology*, 40, 437-454.
- Becton, J., & Schraeder, M. (2004). Participant input into rater selection: Potential effects on the quality and acceptance of ratings in the context of 360-degree feedback. *Public Personnel Management*, 33(1), 23-32.
- Bergman, M. E., Drasgow, F., Donovan, M. A., Henning, J. B., & Juraska, S. E. (2006). Scoring situational judgment tests: Once you get the data, your trouble begins. *International Journal of Selection and Assessment*, 14, 223-235.

- Brutus, S., Fleenor, J. W., & McCauley, C. D. (1999). Demographic and personality predictors of congruence in multi-source ratings. *Journal of Management Development, 18*, 417-435.
- 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, 22*, 395-412.
- Christian, M. S., Edwards, B. D. & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology, 63*, 83-117.
- Conway, J. M., & Huffcutt, A. I. (1997). Psychometric properties of multisource performance ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance, 10*, 331-361.
- Dalessio, A. T. (1998). Using multisource feedback for employee development and personnel decisions. In Smither, J.W. (Eds.), *Performance appraisal: State of the art in practice* (pp.279-330). Jossey-Bass Publishers, San Francisco, CA.
- Dominick, P.C., Reilly, R.R., & Byrne, J.C. (2004). *Individual differences and peer feedback: Personality's impact on behavior change*. Paper presented at the 19th Annual Conference of the Society for Industrial and Organizational Psychology. Chicago, Illinois.
- Foster, C. A. & Law, M. R. F. (2006). How many perspectives provide a compass? Differentiating 360-degree and multi-source feedback. *International Journal of Selection and Assessment, 14*, 288-291.

- Garavan, T., Morley, M., & Flynn, M. (1997). 360 degree feedback: Its role in employee development. *Journal of Management Development*, 16, 134-147.
- Grant, K. L. (2009). *The validation of a situational judgment test to measure leadership behavior*. Unpublished master's thesis. Western Kentucky University, Bowling Green, KY.
- Glenn, S. D., & Horn, T. S. (1993). Psychological and personal predictors of leadership behavior in female soccer athletes. *Journal of Applied Sport Psychology*, 5, 17-34.
- Gould, D., Carson, S., Fifer, A., Lauer, L., & Benham, R. (2009). Social-emotional and life skill development issues characterizing today's high school sport experience. *Journal of Coaching Education*, 2, 1-25.
- Gould, D., Smith, P., White, J., & Chung, Y. (2006). Future directions in coaching life skills: Understanding high school coaches' views and needs. *Athletic Insight: The Online Journal of Sport Psychology*, 8, 1-11.
- Gould, D., & Voelker, D. K. (2010). Youth sport leadership development: Leveraging the sports captaincy experience. *Journal of Sport Psychology in Action*, 1, 1-14.
- Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43-62.
- Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of literature. *Personnel Review*, 37, 426-441.
- Loo, R. (2002). A caveat on using single-item versus multiple-item scales. *Journal of Managerial Psychology*, 17, 68-75.
- Loughead, T. M., Hardy, J., & Eys, M. A. (2006). The nature of athlete leadership. *Journal of Sport Behavior*, 29, 142-158.



- MaCarthy, A., & Garavan, T (2001). 360 feedback processes: Performance improvement and employee career development. *Journal of European Industrial Training*, 25, 5-32.
- Martinek, T., & Hellison, D. (2009). *Youth leadership in sport and physical education*. New York, NY: Palgrave MacMillan.
- McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb III, W. L. (2007). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology*, 60, 63-91.
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A. & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, 86, 730-739.
- McDaniel, M. A., & Nguyen, N. T. (2001). Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment*, 9, 103-113.
- Metcalfe, B. (1998). 360 degree feedback and leadership development. *Professional Forum*, 6, 35-44.
- Motowidlo, S. J., Hanson, M. A., & Crafts, J. L. (1997). Low-fidelity simulations. In D. L. Whetzel & G. R. Wheaton (Eds.) *Applied measurement methods in industrial psychology* (pp. 241-260). Palo Alto, CA: Davies-Black.
- Shoenfelt, E. L. (2009). Situational Assessment of Leadership – Student Assessment (SALSA©): The development and validation of a situational judgment test to assess leadership effectiveness. Unpublished manuscript, Western Kentucky University, Bowling Green, KY.

- Slack, P. (2010). *A situational assessment of student leadership: An evaluation of alternate forms reliability and convergent validity*. Unpublished master's thesis. Western Kentucky University, Bowling Green, KY.
- Smither, J. W., London, M., & Reilly, R. R. (2005). Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings. *Personnel Psychology*, 58, 33-66.
- Smither, J. W., London, M., & Richmond, K. R. (2005). The relationships between leaders' personality and their reactions to and use of multisource feedback: A longitudinal study. *Group and Organization Management*, 30, 181-210.
- Tucker, J. (2011). *An evaluation of the convergent validity of Situational Assessment of Leadership - Student Assessment (SALSA<sup>®</sup>) with multi-source feedback in MBA and Ed.D. in Educational Leadership students*. Unpublished master's thesis. Western Kentucky University, Bowling Green, KY.
- Van Linden, J. A., & Fertman, C. I. (1998). *Youth leadership*. San Francisco, CA: Jossey-Bass.
- Van Velsor, E., & Leslie, J. B. (1991). *Feedback to managers: Vol. 1. A guide to evaluating multi-rater feedback instruments*. Greensboro, NC: Center for Creative Leadership.
- Van Vianen, A., & De Dreu, C. (2001). Personality in teams: Its relationship to social cohesion, task cohesion, and team performance. *European Journal of Work & Organizational Psychology*, 10, 97-120.
- Voelker, D. K., Gould, D., & Crawford, M. J. (2010). The experience of high school sport captains. Manuscript submitted for publication.

- Weekley, J. A., & Ployhart, R. E. (2005). Situational judgment: Antecedents and relationships with performance. *Human Performance, 18*, 81-104.
- Weekley, J. A., Ployhart R. E., & Holtz, B. C. (2006). On the development of situational judgment tests: Issues in item development, scaling, and scoring. In J. A. Weekley, & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 157-182). Mahwah, NJ: Lawrence Erlbaum Associates.
- Weekley, J. A., Ployhart, R. E., & Harold, C. M. (2004). Personality and situational judgment tests across applicant and incumbent settings: An examination of validity, measurement, and subgroup differences. *Human Performance, 17*, 433-461.
- Wood, L., Hassell, A., Whitehouse, A., Bullock, A., & Wall, D. (2006). A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. *Medical Teacher, 28*, 185-191.

## Appendix A

### Description of SALSA© Dimensions

#### **ORGANIZING / PLANNING / VISIONING**

The extent to which the individual systematically arranges his/her own work and resources, as well as that of others, for efficient task accomplishment. The extent to which an individual anticipates and prepares for the future. The extent to which the individual effectively creates an image of the future for the organization and develops the necessary means to achieve that image.

#### **CONSIDERATION / TEAM SKILLS**

The extent to which the individual's actions reflect a consideration for the feelings and needs of others as well as an awareness of the impact and implications of decisions relevant to others inside and outside the organization. The extent to which the individual engages and works in collaboration with other members of the group so that others are involved in the process and the outcome.

#### **PROBLEM SOLVING / INNOVATION**

The extent to which an individual gathers information; understands relevant technical and professional information; effectively analyzes data and information; generates viable options, ideas, and solutions; selects supportable courses of action for problems and situations; uses available resources in new ways; and generates and recognizes creative solutions.

#### **INFLUENCING OTHERS**

The extent to which the individual persuades others to do something or adopt a point of view in order to produce desired results (without creating hostility) and takes action in which the dominant influence is one's own convictions rather than the influence of others' opinions.

#### **COMMUNICATION**

The extent to which the individual effectively conveys both oral and written information. The extent to which the individual effectively responds to questions and challenges.

#### **DRIVE / RESULTS-ORIENTATION**

The extent to which the individual originates and maintains a high activity level, sets high performance standards and persists in achievement, and expresses the desire to advance to higher job levels. The extent to which the individual establishes clear direction, pushes self and others for high quality and results, monitors progress and results, and demonstrates a bias for action.

#### **TOLERANCE FOR STRESS**

The extent to which the individual maintains effectiveness in diverse situations under varying degrees of pressure, opposition, and disappointment.

#### **INTEGRITY / ETHICS**

The extent to which the individual demonstrates consistency between word and deed across situations and circumstances. The extent to which the individual does "the right thing" across situations and circumstances, especially in difficult and challenging situations.

Appendix B

Sample of SALSA© Performance Dimension Rating Form

**CONSIDERATION / TEAM SKILLS**

---

The extent to which the individual's actions reflect a consideration for the feelings and needs of others as well as an awareness of the impact and implications of decisions relevant to others inside and outside the organization. The extent to which the individual engages and works in collaboration with other members of the group so that others are involved in the process and the outcome.

**Includes behaviors such as:**

- **Actively participating in the discussion or activity**
- **Helping to clarify goals**
- **Trying to satisfy group goals**
- **Acknowledging others' feelings**
- **Reinforcing or rewarding others**
- **Accepting and using others' ideas**
- **Welcoming diverging views**
- **Compromising with other group members**
- **Involving others in the discussion/decision/activity**
- **Seeking consensus**
- **Actively seeking contribution from other team members**
- **Diffusing conflict**

**Please rate the targeted athlete on Consideration/Team Skills**

---

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>Very Ineffective</b>	<b>Ineffective</b>	<b>Somewhat Effective</b>	<b>Effective</b>	<b>Very Effective</b>

---

## Appendix C

Correlations between Overall Ratings on each Dimension for each Source  
(Tucker, 2011)

<b>Dimension</b>				
<b><i>OVP</i></b>	Behavioral Average (Self)	Behavioral Average (Peer)	Behavioral Average (Supervisor)	Behavioral Average (Subordinate)
Overall Average (Self)	.73**			
Overall Average (Peer)		.88**		
Overall Average (Supervisor)			.91**	
Overall Average (Subordinate)				.76**
<b><i>CTS</i></b>	Behavioral Average (Self)	Behavioral Average (Peer)	Behavioral Average (Supervisor)	Behavioral Average (Subordinate)
Overall Average (Self)	.77**			
Overall Average (Peer)		.94**		
Overall Average (Supervisor)			.87**	
Overall Average (Subordinate)				.88**
<b><i>PSI</i></b>	Behavioral Average (Self)	Behavioral Average (Peer)	Behavioral Average (Supervisor)	Behavioral Average (Subordinate)
Overall Average (Self)	.63**			
Overall Average (Peer)		.92**		
Overall Average (Supervisor)			.96**	
Overall Average (Subordinate)				.87**
<b><i>INO</i></b>	Behavioral Average (Self)	Behavioral Average (Peer)	Behavioral Average (Supervisor)	Behavioral Average (Subordinate)
Overall	.78**			

Average (Self)				
Overall Average (Peer)		.96**		
Overall Average (Supervisor)			.94**	
Overall Average (Subordinate)				.92**
<b>COM</b>	Behavioral Average (Self)	Behavioral Average (Peer)	Behavioral Average (Supervisor)	Behavioral Average (Subordinate)
Overall Average (Self)	.55*			
Overall Average (Peer)		.90**		
Overall Average (Supervisor)			.91*	
Overall Average (Subordinate)				.87*
<b>DRO</b>	Behavioral Average (Self)	Behavioral Average (Peer)	Behavioral Average (Supervisor)	Behavioral Average (Subordinate)
Overall Average (Self)	.65**			
Overall Average (Peer)		.87**		
Overall Average (Supervisor)			.91**	
Overall Average (Subordinate)				.77**
<b>TFS</b>	Behavioral Average (Self)	Behavioral Average (Peer)	Behavioral Average (Supervisor)	Behavioral Average (Subordinate)
Overall Average (Self)	.88**			
Overall Average (Peer)		.89**		
Overall Average (Supervisor)			.95**	
Overall Average (Subordinate)				.83**

<i>INE</i>	Behavioral Average (Self)	Behavioral Average (Peer)	Behavioral Average (Supervisor)	Behavioral Average (Subordinate)
Overall Average (Self)	.81**			
Overall Average (Peer)		.78**		
Overall Average (Supervisor)			.97**	
Overall Average (Subordinate)				.96**



Appendix D

Correlations by SALSA© dimension

	OPV SALSA	OPV Self	OPV Peer	OPV Coach
OPV SALSA	-			
OPV Self	.256	-		
OPV Peer	.241	.293	-	
OPV Coach	.286	.236	.718**	-
CTS SALSA	.566**	.022	.463*	.411*
CTS Self	-.066	-.006	.164	.097
CTS Peer	-.224	-.044	.527*	.518*
CTS Coach	.115	-.034	.244	.511*
PSI SALSA	.270	.520*	.632**	.434*
PSI Self	.513*	.687**	.432*	.269
PSI Peer	.304	.323	.903**	.711**
PSI Coach	.323	.394	.490*	.712**
INO SALSA	.461*	.406*	.451*	.327
INO Self	.673**	.465*	.413*	.297
INO Peer	-.293	-.158	.406*	.452*
INO Coach	.305	.009	.504*	.509*
COM SALSA	.592**	.270	.416*	.391
COM Self	.385	.242	.417*	.269
COM Peer	.251	.016	.597**	.648**
COM Coach	.317	.165	.437*	.545*
DRO SALSA	.341	-.119	.432*	.372
DRO Self	.026	.371	-.026	.222
DRO Peer	-.335	-.022	.351	.471*
DRO Coach	-.025	-.067	.246	.596**
TFS SALSA	.682**	.240	.269	.292
TFS Self	.042	.465*	.417*	.085
TFS Peer	.183	.119	.331	.378
TFS Coach	.097	-.068	.194	.227
INE SALSA	.543**	.121	.073	.218
INE Self	-.145	.489*	.177	.050
INE Peer	-.462*	-.062	.386	.426*
INE Coach	-.165	-.034	.309	.482*
Total SALSA	.747**	.248	.536*	.483*
OLE Self	.648**	.538*	.431*	.308
OLE Peer	.045	.036	.744**	.712**
OLE Coach	.159	.099	.407*	.652**

\* $p < .05$ , \*\* $p < .01$

	CTS SALSA	CTS Self	CTS Peer	CTS Coach
CTS SALSA	-			
CTS Self	.112	-		
CTS Peer	.067	.292	-	
CTS Coach	.223	.431*	.727**	-
PSI SALSA	.497*	.142	.244	.186
PSI Self	.331	.256	-.138	-.105
PSI Peer	.505*	.142	.548**	.296
PSI Coach	.430*	.037	.204	.288
INO SALSA	.531*	-.191	-.110	-.102
INO Self	.465*	.259	.034	.122
INO Peer	.095	.170	.529*	.311
INO Coach	.367	.093	.097	.122
COM SALSA	.732**	.317	.094	.288
COM Self	.228	.304	.299	.341
COM Peer	.332	.109	.651**	.665**
COM Coach	.312	.233	.471*	.632**
DRO SALSA	.700**	.213	-.062	.018
DRO Self	.116	.131	-.135	.255
DRO Peer	.078	.176	.365	.322
DRO Coach	.451	.023	.466*	.632**
TFS SALSA	.413*	-.060	-.063	.060
TFS Self	.132	.259	-.182	-.282
TFS Peer	.348	-.036	.338	.385
TFS Coach	.513*	.469*	.143	.335
INE SALSA	.562**	-.001	.266	.541*
INE Self	-.034	.456*	.151	.091
INE Peer	.102	.164	.795**	.476*
INE Coach	.037	.179	.732**	.626**
OPV SALSA	.566**	-.066	-.224	.115
OPV Self	.022	-.006	-.044	-.034
OPV Peer	.463*	.164	.527*	.244
OPV Coach	.411*	.097	.518*	.511*
Total SALSA	.912**	.115	.036	.227
OLE Self	.339	.455*	.127	.389
OLE Peer	.410*	.287	.785**	.626**
OLE Coach	.515*	.264	.298	.540*

\* $p < .05$ , \*\* $p < .01$

	PSI SALSA	PSI Self	PSI Peer	PSI Coach
PSI SALSA	-			
PSI Self	.503*	-		
PSI Peer	.641**	.372	-	
PSI Coach	.652**	.367	.510*	-
INO SALSA	.630**	.489*	.382	.292
INO Self	.557**	.659**	.489*	.317
INO Peer	.326	-.262	.565**	.120
INO Coach	.268	.160	.595**	.377
COM SALSA	.647**	.498*	.545**	.569**
COM Self	.454	.342	.537*	.096
COM Peer	.307	.048	.713**	.267
COM Coach	.345	.199	.535*	.603**
DRO SALSA	.362	.343	.295	.458*
DRO Self	.428*	.204	-.001	.467*
DRO Peer	.312	-.120	.223	.156
DRO Coach	.291	-.248	.292	.374
TFS SALSA	.160	.416*	.487*	.163
TFS Self	.476*	.385	.398	.178
TFS Peer	.567**	-.299	.472*	.416*
TFS Coach	.280	-.068	.154	.380
INE SALSA	.131	.036	.268	.126
INE Self	.529*	.296	.219	.266
INE Peer	.309	-.179	.134	.391
INE Coach	.137	-.103	.204	.484*
OPV SALSA	.270	.513*	.304	.304
OPV Self	.520*	.687**	.323	.323
OPV Peer	.632**	.432*	.903**	.903**
OPV Coach	.434*	.269	.711**	.711**
CTS SALSA	.497*	.331	.505*	.430*
CTS Self	.142	.256	.142	.037
CTS Peer	.244	-.138	.548**	.204
CTS Coach	.186	-.105	.296	.288
Total SALSA	.656**	.534*	.585**	.555**
OLE Self	.516*	.725**	.396	.282
OLE Peer	.521*	.108	.806**	.334
OLE Coach	.419	.131	.477*	.650**

\* $p < .05$ , \*\* $p < .01$

	INO SALSA	INO Self	INO Peer	INO Coach
INO SALSA	-			
INO Self	.639**	-		
INO Peer	.099	.033	-	
INO Coach	.269	.330	.467*	-
COM SALSA	.343	.624**	.086	.296
COM Self	.231	.458*	.392	.548
COM Peer	.197	.215	.597	.257*
COM Coach	.074	.326	.212	.569**
DRO SALSA	.485*	.258	-.038	.386
DRO Self	.139	.070	.092	.259
DRO Peer	.315	-.065	.606**	.331
DRO Coach	.245	-.077	.484*	.246
TFS SALSA	.260	.394	.049	.323
TFS Self	.438*	.430*	.206	.401*
TFS Peer	.263	.219	.501*	.129
TFS Coach	.144	.516	.036	.158
INE SALSA	.098	.302	-.029	.117
INE Self	.151	.211	.147	-.121
INE Peer	-.066	-.175	.367	-.071
INE Coach	-.252	-.160	.046	-.135
OPV SALSA	-.293	.673**	-.293	.305
OPV Self	.406*	.465*	-.158	.009
OPV Peer	.451*	.413*	.406*	.504*
OPV Coach	.327	.297	.297	.509*
CTS SALSA	.531*	.465*	.095	.364
CTS Self	-.191	.259	.107	.093
CTS Peer	-.110	.034	.529**	.097
CTS Coach	-.012	.122	.311	.122
PSI SALSA	.630**	.557**	.326	.268
PSI Self	.489*	.659**	-.262	.160
PSI Peer	.382	.489*	.565**	.595**
PSI Coach	.292	.317	.120	.377
Total SALSA	.662**	.659**	.045	.416*
OLE Self	.412*	.767**	-.125	.228
OLE Peer	.295	.280	.769**	.519*
OLE Coach	.290	.193	.405*	.631

\* $p < .05$ , \*\* $p < .01$

	COM SALSA	COM Self	COM Peer	COM Coach
COM SALSA	-			
COM Self	.406*	-		
COM Peer	.541*	.541	-	
COM Coach	.348	.348	.712**	-
DRO SALSA	.521*	.072	.114	.241
DRO Self	.204	.204	.040	.284
DRO Peer	-.120	-.120	.280	.022
DRO Coach	-.248	-.248	.487*	.283
TFS SALSA	.537*	.422*	.463*	.267
TFS Self	.326	.229	-.067	.004
TFS Peer	.363	.237	.394	.242
TFS Coach	.473*	-.026	-.015	.213
INE SALSA	.418*	.418*	.393	.320
INE Self	.222	.222	-.083	-.002
INE Peer	-.088	-.088	.256	.277
INE Coach	-.117	-.117	.275	.443
OPV SALSA	.592**	.385	.251	.317
OPV Self	.270	.242	.016	.165
OPV Peer	.416	.417*	.597**	.437*
OPV Coach	.391	.269	.648**	.545**
CTS SALSA	.331	.331	.332	.312
CTS Self	.256	.256	.109	.233
CTS Peer	-.138	-.138	.651**	.471*
CTS Coach	-.105	-.105	.665**	.632**
PSI SALSA	.647**	.454*	.307	.345
PSI Self	.498*	.342	.048	.199
PSI Peer	.307	.537*	.713**	.535*
PSI Coach	.647**	.096	.267	.603**
INO SALSA	.343	.231	.197	.074
INO Self	.624**	.458*	.215	.326
INO Peer	.086	.392	.597**	.212
INO Coach	.296	.548**	.527*	.569**
Total SALSA	.853**	.405	.369	.390
OLE Self	.551**	.623**	.321	.420
OLE Peer	.347	.584**	.864**	.554**
OLE Coach	.437	.225	.600**	.753**

\* $p < .05$ , \*\* $p < .01$

	DRO SALSA	DRO Self	DRO Peer	DRO Coach
DRO SALSA	-			
DRO Self	.092	-		
DRO Peer	.224	.380	-	
DRO Coach	.302	.397	.675**	-
TFS SALSA	.240	-.123	-.306	.029
TFS Self	.287	.196	.272	-.112
TFS Peer	-.013	.176	.203	.505*
TFS Coach	.610**	.203	.252	.486*
INE SALSA	.000	.104	-.129	.469*
INE Self	.156	.286	.203	.115
INE Peer	.155	.000	.462*	.521*
INE Coach	.084	-.039	.181	.371
OPV SALSA	.341	.026	-.355	-.025
OPV Self	-.119	.371	-.022	-.067
OPV Peer	.432*	-.026	.351	.246
OPV Coach	.372	.222	.471*	.596**
CTS SALSA	.700**	.116	.078	.451*
CTS Self	.213	.131	.176	.023
CTS Peer	-.062	-.135	.365	.466*
CTS Coach	.018	.255	.322	.632**
PSI SALSA	.362	.428	.312	.291
PSI Self	.343	.204	-.120	-.248
PSI Peer	.295	-.001	.292	.292
PSI Coach	.458*	.467*	.374	.374
INO SALSA	.485	.139	.315	.245
INO Self	.258	.070	-.065	-.077
INO Peer	-.038	.092	.606**	.484*
INO Coach	.386	.259	.331	.246
COM SALSA	.521*	.313	-.027	.268
COM Self	.027	.236	.253	.175
COM Peer	.280	.040	.114	.487*
COM Coach	.022	.284	.241	.283
Total SALSA	.710**	.203	.005	.372
OLE Self	.184	.315	-.009	-.017
OLE Peer	.247	.109	.564**	.606**
OLE Coach	.561	.515	.444*	.662**

\* $p < .05$ , \*\* $p < .01$

	TFS SALSA	TFS Self	TFS Peer	TFS Coach
TFS SALSA	-			
TFS Self	.130	-		
TFS Peer	.094	.118	-	
TFS Coach	-.078	.307	.394	-
INE SALSA	.428*	-.285	.387	.256
INE Self	.091	.567**	.253	.434*
INE Peer	-.370	-.134	.246	.318
INE Coach	-.258	-.368	.177	.258
OPV SALSA	.982**	.042	.183	.097
OPV Self	.240	.465*	.119	-.068
OPV Peer	.269	.417*	.331	.196
OPV Coach	.292	.085	.378	.227
CTS SALSA	.912**	.132	.348	.513*
CTS Self	-.060	.259	-.036	.469*
CTS Peer	-.063	-.182	.338	.143
CTS Coach	.060	-.282	.385	.335
PSI SALSA	.160	.476*	.567**	.280
PSI Self	.416*	.385	-.299	-.068
PSI Peer	.585**	.472*	.472*	.154
PSI Coach	.163	.416*	.416*	.380
INO SALSA	.260	.438*	.263	.144
INO Self	.394	.430*	.219	.156
INO Peer	.049	.206	.501*	.036
INO Coach	.323	.401*	.129	.158
COM SALSA	.537*	.326	.363	.473*
COM Self	.422	.229	.237	-.026
COM Peer	.463	-.067	.394	-.015
COM Coach	.267	.004	.242	.213
DRO SALSA	.240	.287	-.013	.610**
DRO Self	-.123	.196	.176	.203
DRO Peer	-.306	.272	.203	.252
DRO Coach	.029	-.112	.505*	.486*
Total SALSA	.584**	.085	.274	.465*
OLE Self	.336	.268	.137	.151
OLE Peer	.243	.114	.437*	.171
OLE Coach	.525*	.224	.290	.504*

\* $p < .05$ , \*\* $p < .01$

	INE SALSA	INE Self	INE Peer	INE Coach
INE SALSA	-			
INE Self	-.105	-		
INE Peer	.014	.246	-	
INE Coach	.139	.057	.832**	-
OPV SALSA	.543**	-.145	-.462*	-.165
OPV Self	.121	.489*	.062	-.309
OPV Peer	.073	.177	.386	.309
OPV Coach	.218	.050	.426*	.482*
CTS SALSA	.562**	-.034	.037	.037
CTS Self	-.001	.456*	.179	.179
CTS Peer	.266	.151	.732**	.732**
CTS Coach	.541*	.091	.626**	.626**
PSI SALSA	.131	.529*	.309	.137
PSI Self	.036	.296	-.179	-.103
PSI Peer	.268	.219	.314	.204
PSI Coach	.126	.266	.391	.484*
INO SALSA	.098	.151	-.066	-.252
INO Self	.302	.211	-.175	-.160
INO Peer	-.029	.147	.367	.046
INO Coach	.117	-.121	-.071	-.135
COM SALSA	.481*	.400*	.062	.063
COM Self	.418	.222	-.088	-.117
COM Peer	.393	-.083	.256	.275
COM Coach	.320	.277	.277	.443*
DRO SALSA	.000	.156	.155	.084
DRO Self	.104	.286	-.039	-.039
DRO Peer	-.129	.203	.181	.181
DRO Coach	.469*	.115	.371	.371
TFS SALSA	.428*	.091	-.370	-.258
TFS Self	-.285	.567**	-.134	.368
TFS Peer	.387	.253	.246	.177
TFS Coach	.256	.434*	.318	.258
Total SALSA	.530*	.172	.002	-.001
OLE Self	.381	.269	-.180	.002
OLE Peer	.291	.154	.532*	.354
OLE Coach	.219	.181	.330	.273

\* $p < .05$ , \*\* $p < .01$



