SELF-EFFICACY AND PERFORMANCE IN VOLLEYBALL REFEREES

By

Benjamin D. Spencer

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Kinesiology - Master of Science

2015

ABSTRACT

SELF-EFFICACY AND PERFORMANCE IN VOLLEYBALL REFEREES

By

Benjamin D. Spencer

Sport officials are an under-researched subpopulation in the sport psychology literature. Particularly little is known about psychological factors that may predict officiating performance. Feltz and Guillen (2011) suggested that self-efficacy may influence performance in the refereeing context, as it does in many others. Myers, Feltz, Guillen, and Dithurbide (2012) indicated that referee self-efficacy is composed of four dimensions: Game Knowledge, Decision-Making, Pressure, and Communication. The current study sought to evaluate the relationship of these various dimensions to performance in several aspects of officiating. A secondary purpose was to evaluate proposed sources of referee efficacy as predictors of referee efficacy dimensions and performance. Volleyball referees (N = 76) who were candidates for USA Volleyball (USAV) National or Junior National badges completed a survey which measured experience as an official, experience playing and coaching volleyball, referee self-efficacy, and sources of referee self-efficacy. Following administration of the survey, participants completed the USAV referee performance evaluation protocol. No relationship was found between self-efficacy and performance in high-level volleyball officials. Little was found relating referees' level of experience and performance in their evaluations, and few connections were identified between previously established sources of referee confidence and dimensions of referee self-efficacy. These null findings may be due to lack of variance in ability and confidence on the part of the referees, or produced by an evaluation system which is designed to teach candidates, and pass most of them, rather than explicitly evaluate their performance.

For my wife, Meghan, who has made so many sacrifices so that I may continue to learn, and without whom I never would have finished Figure 2;

And for my father, who set me on this path long ago, though I doubt either of us would have guessed it would look like this.

ACKNOWLEDGMENTS

First, I would like to acknowledge the support and guidance of my advisor, Dr. Deborah Feltz, in the development of this thesis and during my two years at Michigan State. I must also thank the members of my committee, Dr. Al Smith and Dr. Francisco Villarruel, for their patience and their helpful feedback.

The support of my class- and lab-mates was priceless to me as I worked through this project and my time in the master's program. Faith Lincoln, Tayo Moss, Steve Samendinger, Alison Ede, Emery Max, and Christel Beverly: I could not have done this without you. Special mention must be given to Anthony Delli Paoli and Dr. Andy Driska, for helping me get moving when I was stuck. There are more names I should mention, but I haven't enough space. I never had a negative experience with anyone in the Kinesiology department, which is astonishing to think about.

As always, I am nothing without my family. The love and support of my parents, Dave and Pat; my brother, Nick; and my wife, Meghan are, as ever, the foundation upon which my life is built. None of this would be possible without Meghan's herculean efforts to keep a roof over our heads and food on the table, after I dragged her halfway across the country to be unemployed.

Finally, I must thank the wonderful people at USA Volleyball, in particular Michelle Prater and Michael O'Connor. I felt welcomed and was accommodated during this project beyond my wildest expectations.

To all of you, thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Nature of the Problem

Research in sport psychology is traditionally focused on coaches and athletes. Referees are a sub-population with important roles in sports, but are largely ignored in the current literature. McInman (1997) analyzed four major sports psychology journals over a 10-year period, and found that only 1.12% of articles addressed or involved officiating. The frequency with which players, coaches, and spectators blame officials for their teams' competitive failures implies that an increased understanding of the factors underlying referee performance would be productive for nearly everyone involved in sports at any level.

The most important predictor of referee performance is experience. Years of experience officiating, number of matches officiated, and hours of practice are positively related to skill (Catteeuw, Helsen, Gilis, & Wagemanns, 2011), although officiating is somewhat unique in that deliberate practice is difficult (MacMahon, Helsen, Starks, & Weston, 2007). The best way to practice is simply to referee more matches. Experience playing or watching the game may also be helpful (Pizzera & Raab, 2012), and in some sports may be able to substitute for refereeing experience early in an official's career (Dosseville, Labord, & Raab, 2011) or even be more valuable than extensive experience as a referee. This means that for some sports, officials should cease competitive play and specialize as a referee early, while in others fast-tracking former high-level competitors may be the most productive avenue for producing high level referees (Pizzera & Raab, 2011). Volleyball has not been investigated as to whether "early specialists" or fast-tracked competitors make the best officials, but it appears that former competitors tend to

make better referees in very specific, technical sports such as judo and trampolining, while more traditional sports (like volleyball) tend to favor early specialists.

Defining and measuring referee performance can be challenging. While athletes and coaches can often be evaluated on objective, observable outcomes—wins and losses, times, scores, batting averages, etc—referees must be evaluated on frequent and often subjective interpretations of incidents during play. Rather than producing a quantifiable end result, sports officials facilitate a safe and fair contest for the participants to a greater or lesser degree. This broad directive results in two general approaches to assessing or studying referee performance: an incident-by-incident, right-or-wrong objective analysis of specific decisions during the course of a contest; and a more holistic appraisal of more subjective elements of officiating such as match control and personnel management. Reviewing this literature is complicated by the reality that specific officiating tasks vary from sport to sport, however it is reasonable to assume that some commonalities can be inferred.

The first approach lends itself well to laboratory-setting study of perceptual and decision-making processes. MacMahon et al. (2007) showed that referees were more accurate than soccer players in judging videotaped challenges, implying that incidental decision-making is a skill that referees develop through experience and practice. Catteuw et al. (2011) went a step further, finding that soccer referees and assistant referees (linesmen) each perform better on tasks specific to their role. The oft-studied officiating biases are apparent in individual decisions. Multiple studies have attributed home advantage at least in part to the home crowd's influence upon officials. This effect has been identified across multiple sports, leagues, and countries, including the English Premier League (Boyko, Boyko, & Boyko, 2007) and the NBA (Lehman & Reifman, 2001). Other studies have noted racial biases (e.g. Wagner-Egger, Gygax, &

Ribordy, 2012), differences when officiating men versus women (Souchon et al. 2004, 2009a, 2009b, 2010), and even a bias against taller athletes (van Quaquebeke & Leissner, 2010).

The second, more holistic approach to evaluation is more difficult and less common in research. Game management is sometimes quantified as use of the oft-derided "make-up call," which can be conceptualized as another form of bias. NCAA basketball referees are more likely to call fouls on teams which so far have fewer fouls than their opponents, and that tendency increases in strength as the gap widens (Anderson & Pierce, 2009). Lopez and Snyder (2013) noted that NHL referees will issue make-up calls over the course of a game to even out the number of penalties assessed to each team, with the aim of achieving perceptions of balance and fairness, a practice they termed "biased impartiality." Soccer referees are less likely to award a penalty to a team that has already received one, and more likely to award a penalty to a team who has previously had one given against them (Plessner & Betch, 2001). Referees manage the flow of a competition by adjusting their decision-making depending on the context. For example, soccer officials who view recorded incidents in the context of a match award fewer yellow cards than when watching incidents in random sequence, and are more likely to award yellow cards when told that an incident is late in the match rather than early (Unkelbach & Memmert, 2008).

Despite the difficulties and complexities of evaluating referee performance, leagues or officials' organizations must assess their referees to determine readiness for promotion and level for assignment, and to assure quality for the competition. Accordingly, they must either find assessment tools in the literature or create their own. Anshel (1995) developed the Behaviorally Anchored Rating Scale for Basketball Referees (BARS-BR) in an attempt to provide an objective measure of officiating performance. Unfortunately, application of the measure is, as the name suggests, limited to basketball referees. The Systematic Observation of Referees'

Behavior (Trudel, Cote & Sylvestre, 1996) quantifies the time that referees spend in various activities (e.g. monitoring without interaction, intervening verbally or with gestures), but does not provide a judgment of performance quality.

Leagues and referee organizations attempt to evaluate their officials through observation, approximately objective scoring and feedback from veteran and expert referees. The population of interest in the present study is USA Volleyball (USAV) referees. USAV uses their own rating sheets to evaluate performance in candidates for Junior National or National certification. Experienced, high level referees rate candidates on multiple dimensions of performance, including judgment (e.g., consistency), mechanics/signals (e.g., scanning the court before beckoning for the serve), positioning/focus (e.g., watching each ball contact), match control (e.g., warm-up administration), communication with match participants (e.g., demeanor and approachability), communication with officiating team (e.g., interactions with line judges, scorekeeper, and second referee), and professionalism (e.g., appearance). These ratings serve as the performance outcome measure in the present study. Notably, this instrument attempts to score candidates on both the subjective, match control tasks and the objective, single-incident tasks.

Guillen and Feltz (2011) proposed a conceptual model of referee efficacy, which they defined as the extent to which referees believe they have the capacity to perform successfully in their job. Based on Bandura's (1977, 1997) theory of self-efficacy, referee self-efficacy was suggested to include six dimensions: game knowledge, decision-making skills, psychological skills, strategic skills, communication/control of the game, and physical fitness. Guillen and Feltz proposed four sources of referee self-efficacy based on the Sources of Sport Confidence Scale (Vealey, Hayashi, Garner-Holman, & Giacobbi, 1998). These sources include mastery

experience, significant others, physical and mental preparation, and partner qualifications. Finally, the authors speculated that self-efficacy would influence referee behavior, satisfaction, stress, and performance, as well as athlete rule violations and coach behavior.

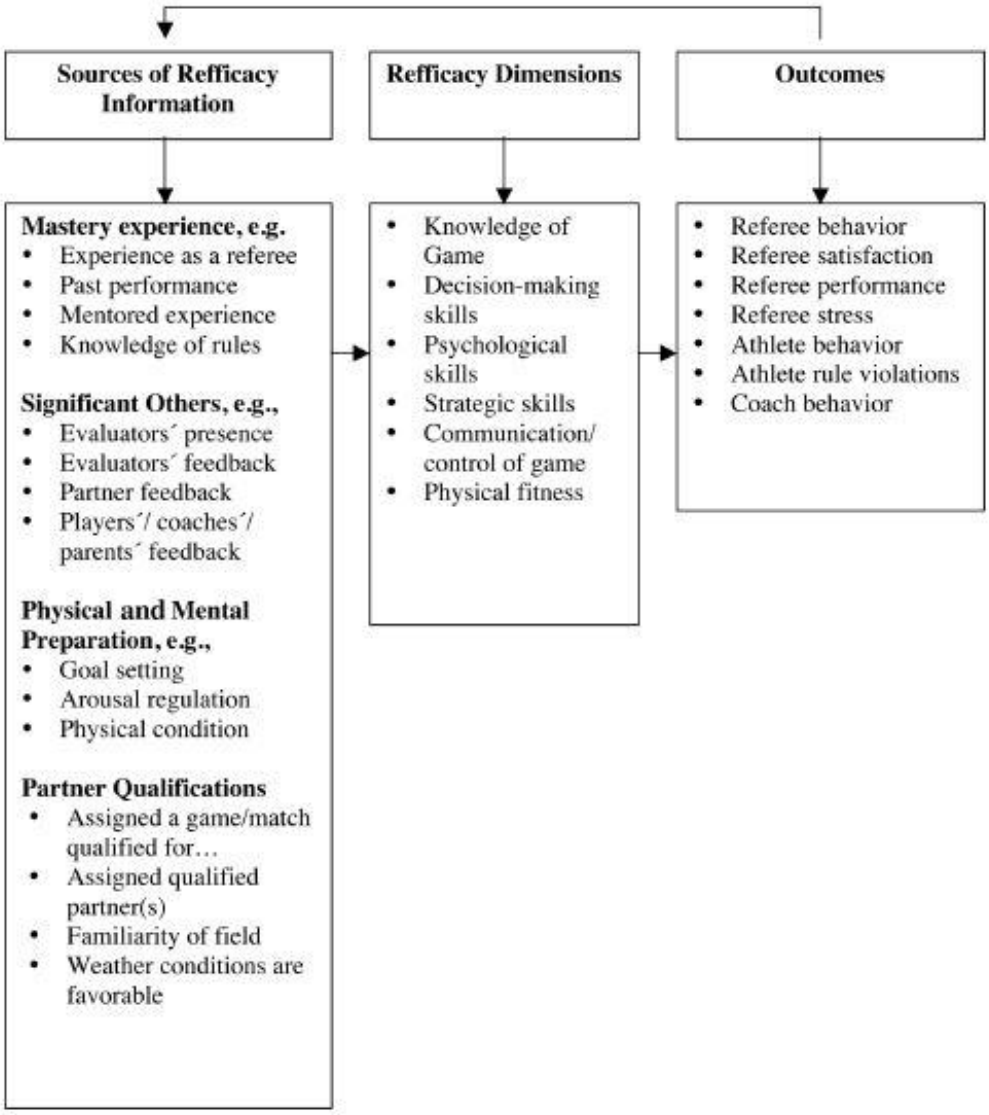

Figure 1: Conceptual Model of Referee Self-Efficacy (Guillen & Feltz, 2011)

Myers, Feltz, Guillen, and Dithurbide (2012) expanded on that theoretical framework to develop a specialized measure for evaluating self-efficacy in referees. The Referee Self-Efficacy Scale (REFS) consists of 39 items which measure four factors of referee self-efficacy: game knowledge, decision making, pressure, and communication. Game knowledge was defined as

the confidence that referees have in their knowledge of their sport, including rules, officiating mechanics, and basic game strategy. Decision making was defined as the confidence that referees have in their ability to quickly and firmly make decisions during competition. Pressure was defined as the confidence that referees have in their ability to be uninfluenced by pressure from players, spectators, and coaches. Communication was defined as the confidence that referees have in their ability to communicate effectively with other referees, coaches, players, and auxiliary personnel. Myers et al. demonstrated factorial validity for the four-dimension REFS with a large sample of referees representing 15 different team sports from the US and Spain. In addition, the authors showed support for the sources of referee self-efficacy as significant predictors of the four dimensions of REFS.

While Myers et al. (2012) provided preliminary support for the first part of the Guillen and Feltz (2011) model, no research has examined the outcomes of self-efficacy in referees—such as performance. Myers et al. suggested that such an investigation could make an important contribution to the literature, especially if examined simultaneously with proposed sources of referee efficacy (e.g., exploring dimensions of referee efficacy as mediators).

Purpose of the Study

The purpose of the present study was to examine the predictive strength of referee self-efficacy on officiating performance in volleyball referees. This research represents a clear gap in the existing literature: self-efficacy and performance have never been studied together within the referee population. Additionally, this research sought to explore the mediating role of the dimensions of referee efficacy between referee self-efficacy sources and performance in volleyball referees.

Proposed Model

The study tested the hypothesized relationships in the proposed model (see Figure 2).

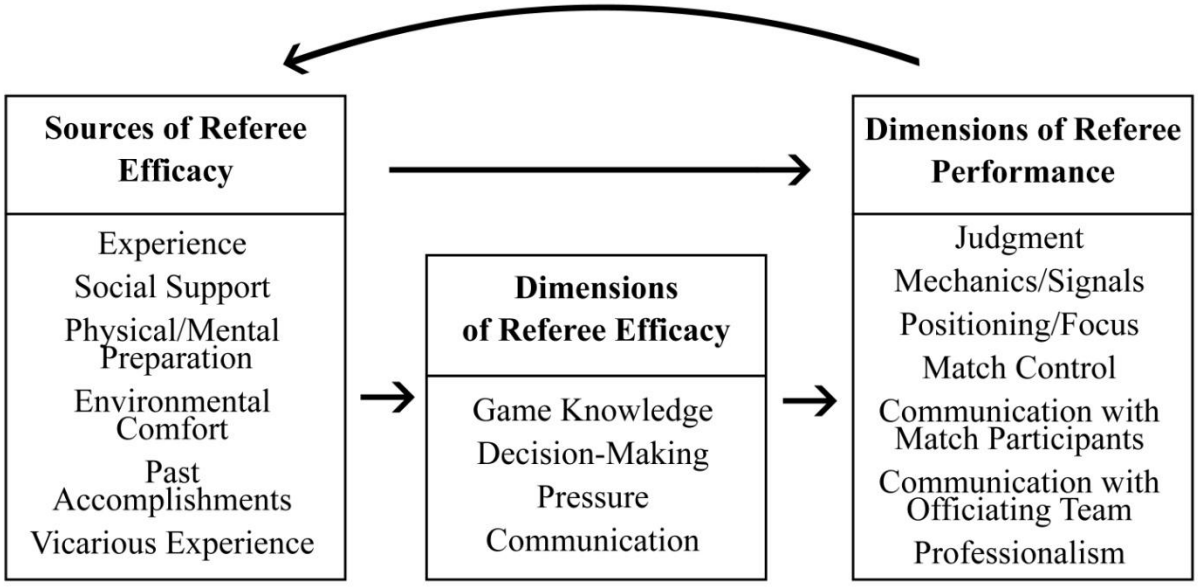

Figure 2: The Hypothesized Model.

The proposed sources of referee efficacy are hypothesized to predict the dimensions of referee efficacy identified in Myers et al. (2012) as well as the various dimensions of performance on which volleyball referees

Delimitations

1. The population was delimited to USAV volleyball referees who are candidates to become National or Junior National officials.
2. The referee self-efficacy of the officials was measured by the Referee Self-Efficacy Scale (REFS) (Myers et al., 2012).
3. The referees' sources of referee self-efficacy was measured by the Sport Officials Self-Rating Scale (Guillen & Feltz, 2011).
4. Referee performance was measured by the USAV First Referee Rating Sheet.

Limitations

1. Myers et al. (2012) suggested a minimum sample size of 300 for use of the REFS.
2. Not all possible determinants of referee performance were measured in this study.

Definitions

1. *Referee efficacy*: The extent to which a referee believes that he or she has the ability to successfully officiate a competition (Myers et al., 2012).
2. *Referee, official, sport official*: These are used interchangeably throughout this manuscript, and can also be assumed to include other sport-specific titles such as umpire, judge, or technical official, which refer to an authority figure responsible for presiding over a sport competition and enforcing the rules from a neutral point of view.
3. *Self-efficacy*: Academically, self-efficacy is beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments (Bandura, 1997).

CHAPTER 2

REVIEW OF LITERATURE

Self-Efficacy Theory

The theory of self-efficacy (Bandura, 1977, 1997, 2001) was developed within the framework of social cognitive theory. Social cognitive theory views behavior from an agentic perspective (Bandura, 2001); that is, individuals use forethought, self-reflection, and self-regulation to influence their own functioning rather than passively react to their environment. These agentic behaviors interact along with personal factors and environmental conditions to determine motivation and behavior (Bandura, 1986, 1997). Self-referent thought mediates the relationship between knowledge and behavior, and as such individuals' evaluations of their capabilities and self-perceptions of efficacy affect their motivation and performance. Self-efficacy is, according to Bandura, the most influential form of self-belief.

Bandura defines self-efficacy as a judgment about one's capability to successfully perform a task at given levels. Efficacy involves not only knowing what behavior is appropriate for a situation, but also organizing cognitive, social, and behavioral strategies and skills to produce the correct action. Thus, judgments of efficacy are not based on skills alone, but instead on what an individual can do with the skills they possess. Efficacy expectations, according to Bandura, should not be confused with outcome expectations. While efficacy expectations reflect one's belief in their capability, outcome expectations illustrate beliefs about the likely consequence of a behavior, such as recognition, rejection, rewards, or punishment. The critical distinction is that an individual might believe that certain behavior will result in a desired outcome, but their execution of that behavior in the end will be more dependent upon their beliefs in their capability than on their beliefs in regard to outcome.

Self-efficacy beliefs vary and are measured across three dimensions: level, strength, and generality (Bandura, 1997). An individual's level of self-efficacy reflects their expected performance at a given level of difficulty. Volleyball referees with disparate levels of self-efficacy for their ability to judge "ballhandling" infractions, for example, might judge how many of such situations they could assess correctly in a given number of opportunities (e.g., 1 out of 10, 5 out of 10, 10 out of 10). Strength of self-efficacy beliefs is the individual's degree of certainty that they can attain a level of performance. Two referees might both believe that they can correctly call 10 out of 10 possible ballhandling infractions, but one might be much more certain about their ability to do so than the other. Generality refers to the degree to which an individual considers themselves efficacious in numerous tasks or domains, or to transfer efficacy judgments from one task to another. A referee with a large degree of generality in their efficacy beliefs might be able to transfer their efficacy for calling ballhandling infractions into other decisions, or even other sports. This is important, because self-efficacy beliefs are not a global trait for an individual; rather they are specific to distinct domains of functioning and even specific aspects of a given domain. For example, a referee might have high levels of efficacy for judging infractions, but low efficacy for communication with coaches and players.

Feltz, Short, and Sullivan (2008) identify several distinct types of efficacy beliefs relevant to sport, many of which may be important to the referee subpopulation. The most straightforward type is perhaps *task self-efficacy*, beliefs about performing a particular task with graded levels of difficulty. The previously mentioned ballhandling infractions are an example of task self-efficacy. *Ameliorative* or *coping efficacy* relates to an individual's belief in their ability to manage perceived threats. For referees, these threats might appear in the forms of stress from pressure applied by match participants, coaches, and specatators, distractions or difficulties

presented by the tournament environment, anxiety resulting from the evaluation process, or a need to recover from mistakes. *Performance efficacy* is an individual's belief in their ability to successfully complete a task at a specific time or in specific context, rather than in general. An official might feel good about their ability to judge ballhandling typically, but for whatever reason not feel so confident on the day of a tournament. *Collective efficacy* describes a group's shared beliefs in its ability to perform at a given level. For officials in many sports, including volleyball, communication and teamwork among a group of referees is crucial for successful performance. Other types of efficacy beliefs less pertinent to the current study include *self-regulatory efficacy*, *learning efficacy*, and *competitive efficacy*.

Sources of Efficacy Information

Bandura (1997) proposes four sources of efficacy information: past performance, vicarious experience, verbal persuasion, and physiological/emotional states. These sources impact individuals' behavior patterns (such as persistence, effort, and choice) and thought patterns (such as goals and attributions) indirectly via their influence on efficacy expectations. A person may draw on one or more of these sources to form their efficacy beliefs, and the salience of the various sources may differ across individuals or tasks.

Sources of past performance efficacy information are dependent on an individual's mastery experiences and accomplishments. Usually, efficacy expectations will increase with past performance that the individual perceives as successful, and lower with perceived unsuccessful experiences. The efficacy value of performance experiences is also impacted by task difficulty, temporal pattern of success and failure, effort expended, and guidance received (Bandura, 1982). Success with minimal effort on tasks that are considered difficult implies high ability and enhances efficacy beliefs. Additionally, efficacy beliefs are impacted by the individual's

conception of an ability as an inherent aptitude as opposed to an acquired skill (Bandura, 1997). Bandura considers past performance to be the most influential source of efficacy information.

Vicarious experience involves observation and comparison of oneself to others performing a task. Watching or simply visualizing others succeed can increase self-efficacy for a task, while seeing others fail to perform successfully can lower efficacy expectations. This modeling effect is strongest when the observer is similar to the model. Efficacy beliefs can also be affected by social comparison with others. Weinberg, Gould, and Jackson (1979) showed that self-efficacy could be manipulated by portraying a competitor as competent or incompetent. Self-modeling, or observation of one's own past performances, can also enhance self-efficacy (Dowrick & Dove, 1980). This may extend to imaginal experiences in which an individual visualizes themselves or others behaving successfully or successfully, though Maddux (1995) argues that imaginal experiences should be considered a distinct source of efficacy information, rather than being included in vicarious experience. While vicarious experiences are believed to be a weaker source of self-efficacy than past performance, they remain influential. When an individual is convinced through vicarious experience that they lack efficacy for a task, they may act in ways that confirm that notion.

Verbal persuasion comes in myriad forms, including others' expectations and evaluative feedback as well as self-talk, positive imagery, and other cognitive strategies. According to Bandura, it is easier to lower efficacy beliefs through criticism than it is to raise them through positive feedback. The extent of verbal persuasion's influence on efficacy beliefs is dependent on the source's credibility and trustworthiness to the individual receiving the feedback, as well as the realistic nature of the information imparted. Individuals can regulate their thought processes

to convince themselves that they can perform at a certain level through self-talk (Feltz et al., 2008).

Physiological and emotional states which influence self-efficacy include autonomic arousal, pain, fear, fatigue, and stress, among others. In general, positive states and emotions, or the lack of negative states, will increase efficacy expectations, while negative physiological states will lower efficacy expectations. Individuals evaluate and interpret these states differently; for example, one person might interpret the feeling of butterflies in their stomach as anxiety and an indication that they are not prepared for a coming experience, while another individual might associate that same feeling with readiness to perform. Feltz et al. (2008) separated physiological information and emotional states into independent categories for sources of self-efficacy in sport contexts because they relate to different aspects of performance. With regard to the current study: physiological information is a more important efficacy source for physical tasks as opposed to non-physically demanding performances, and officiating a volleyball match is not inherently a physically demanding task, so affective states may be of more interest.

Self-Efficacy and Performance

Bandura (1986) suggests that self-efficacy contributes to behavior in multiple ways. Self-efficacy beliefs influence how people behave, their emotional reactions to events, and their thought patterns in various situations. People tend to avoid situations which they do not believe they are capable of succeeding in, and their level of self-efficacy helps determine the degree of effort and persistence they will show when facing failure. Emotional reactions and thoughts are affected by one's efficacy in regards to effort, stress, and attentional demand. Those with high self-efficacy can focus on tasks at hand and produce more effort in comparison to people with low self-efficacy, who may be anxious and divert attention from possible solutions. Notably,

efficacy judgments only a major determinant of behavior when requisite skills and proper incentives are present; a referee without any experience is unlikely to be successful if thrown into a high-pressure situation, regardless of their belief in their capability.

Self-efficacy affects people in innumerable domains, from education, to politics, to sport performance. High self-efficacy not only yields better performance, but efficacious individuals are less afraid to set challenging goals and persevere through failure (Feltz et al., 2008). Though complacency can set in during periods of success, a meta-analysis conducted by Moritz, Feltz, Fahrbach, and Mack (2000) showed self-efficacy to have an average correlation of performance of .38 when individuals have incentive to act on their efficacy beliefs, when they possess the requisite skills, when the nature of the task is clear, and when the measure used for obtaining efficacy and performance data is unambiguous. Thus, self-efficacy has a moderately positive effect on subsequent sport performance.

Self-Efficacy in Sport Officials

Only recently have researchers begun studying self-efficacy in the sport officiating context. Because the officiating role is distinct from others in sports in its non-competitive nature, its unique pressures, and its tasks which are very different than those of players or coaches, models used to study efficacy beliefs in other populations (such as players and coaches) are not suitable for application to referees. Guillen and Feltz (2011) offered a conceptual framework for referee efficacy, which included referee-specific sources of efficacy information as well as effects or outcomes of efficacy in referees. This model was developed as a result of a focus group of Midwestern soccer referees with various levels of experience. Participants were asked to identify what they believed to be the key areas of referee efficacy needed to perform their job as an official, the sources of their efficacy, and the influence of those efficacy beliefs on

their and others' behavior. Based on these discussions, referee self-efficacy was suggested to include six dimensions: game knowledge, decision-making skills, psychological skills, strategic skills, communication/control of the game, and physical fitness. *Game knowledge* reflects the referee's critical need for adequate knowledge of physical, technical, and tactical aspects of the sport they officiate. *Decision-making skills* refers to making critical decisions accurately during competition, as well as being firm in one's decisions. *Psychological skills* references officials' need to focus attention and concentration, recover from bad calls, and demonstrate poise. *Strategic skills* include proper positioning on the area of play, consistent proper mechanics and signals, and anticipation of game actions. *Communication/control of game* relates to communication with players, coaches, and other officials, as well as resolving disputes and making necessary adjustments in their behavior and decisions to maintain control of the game. *Physical fitness* was deemed important in sports where referees engage in a lot of physical exercise, as good fitness is a requirement to stay with the play.

Vealey, Hayashi, Garner-Holman, and Giacobbi (1998) identified nine sources of sport confidence in athletes: mastery, demonstration of ability, physical and mental preparation, physical self-presentation, social support, coaches' leadership, vicarious experience, environmental comfort, and situational favorableness. Sport confidence, as defined by Vealey et al. as "the degree of certainty individuals possess about their ability to be successful in sport," fits with self-efficacy because both describe what people perceive they can do, rather than what they have or what they are (Feltz et al., 2008). Using these sources of confidence, combined with feedback from the focus group and Bandura's (1997) sources of self-efficacy, Guillen and Feltz (2011) proposed four major sources of self-efficacy for referees. *Mastery experience* involves years of referee experience, past performance, mentored experience, and knowledge of the rules,

and as with the performance accomplishments source of self-efficacy in Bandura's (1997) theory, was expected to be the strongest source of efficacy. *Significant others* refers to the referee's perceived level of support or non-support from players, coaches, spectators, peers and partners, and evaluators or administrators, as well as social comparison with other referees. *Physical and mental preparation* relates to goal-setting, arousal regulation, self-talk, visualization of good performance, and readiness for maximum effort. The final source, *partner qualifications,* reflects officials' preference for working at certain levels, in specific facilities, and with qualified, able, familiar partners.

Myers, Feltz, Guillen, and Dithurbide (2012) developed an instrument to measure and quantify the framework proposed by Guillen and Feltz (2011). The Referee Self-Efficacy Scale (REFS) was developed based on themes from a focus group of soccer referees as well as relevant conceptual and measurement literature. It is composed of 39 items which measure four factors of referee self-efficacy: game knowledge, decision making, pressure, and communication. After initial testing on several hundred soccer referees in the United States and Spain, two of the dimensions proposed in Guillen and Feltz (2011), psychological skills and control of the game, were collapsed into the pressure dimension on the REFS. This decision was supported by a single-group exploratory structural equation model, in which a five-factor solution did not significantly improve upon the four factor solution accepted as the final model. Myers et al. showed validity for the REFS across two countries (United States and Spain), levels of competition (youth, high school, and elite, which included collegiate, professional, and international referees), team gender, and sports (soccer and basketball). In addition, the authors showed support for the most of the sources of referee self-efficacy as significant predictors of the four dimensions of referee efficacy. While social support did not predict any of the four

dimensions, physical/mental preparation predicted all four, environmental comfort and situational favorableness predicted decision making, pressure, and communication, and past accomplishments and vicarious experience predicted game knowledge, decision making, and pressure.

To my knowledge, self-efficacy's effect on performance has not been studied in referees. This represents a gap in the literature which the current study sought to fill.

CHAPTER 3

METHOD

Participants

The participants in this study were American volleyball referees at the National and Junior National levels. The total number of registered volleyball referees at these levels in the United States is about 700. Participants for this study were 76 referees: 30 applying for National certification, 46 applying for Junior National certification, which represents about 10% of the targeted population. Of the 60 participants who reported their gender, 40 (66.7%) were male, 20 (33.3%) female. 59 (77.6%) of participants described themselves white/Caucasian, 5 (6.6%) chose black/African-American, 5 (6.6%) reported themselves as Hispanic, 2 (2.6%) were Asian/pacific islander, 3 (3.9%) chose "other", and 2 (2.6%) did not report their race. Among candidates who reported their age (*N*=19), ages ranged from 23 to 63, with a mean of 44.

Instruments

Sports Officials Self-Rating Scale (A modified version of the Sources of Sport Confidence Scale; See Appendix B) (Vealey, Hayashi, Garner-Holman, & Giacobbi, 1998*)*: The Sports Officials Self-Rating Scale (Guillen & Feltz, 2011) was used to evaluate sources of self-efficacy in the participating referees. Officials are asked to indicate how important various events are in giving them confidence in officiating their sport. The measure has 25 items on a 7-point scale. The stem for all items is "I gain confidence in officiating when I…." The 7-point scale is as follows: 1=Not at all important, 2=Not very important, 3=Slightly important, 4=Of average importance, 5=Very important, 6=Extremely important, 7=Of highest importance. Each item relates to one of six sources of self-efficacy: social support (e.g., *Get positive feedback from other officials*), physical or mental preparation (e.g., *Keep my focus on the game*), environmental

comfort (e.g., *Officiate in a venue I like*), situational favorableness (e.g., *Am familiar with officials I will officiate with*), past accomplishments (e.g., *Performed well in previous contests*), and vicarious experience (e.g., *See successful officiating by other officials in my sport*). One item, "*Weather conditions are favorable,"* was changed to "*Venue conditions are favorable"* to improve its relevance to indoor volleyball specifically. Included with the Sports Officials Self-Rating Scale was a short questionnaire that collects demographics and background information about the participant. This information includes age, gender, USAV region, years of experience playing, coaching, and officiating volleyball, highest level playing, coaching, and officiating volleyball, other sports refereed (if any), number of matches officiated in the past year, and number of training sessions or clinics attended in the past year.

Referee Self-Efficacy Scale (REFS; See Appendix C): The REFS (Myers, Feltz, Guillen, & Dithurbide, 2012) was used to evaluate self-efficacy in the participating referees. The REFS has 39 items on a 5-category rating scale. The stem for all items is "In relation to the primary sport(s) that you referee, how confident are you in your ability to…." Each item relates to one or

Table 1. Adjusted items in the Referee Self-Efficacy Scale

Item #	Original item	Adjusted item
2	Know when and how to call more fouls/penalties to control the flow of the game	Know when and how to call more or fewer faults/infractions to control the flow of the game
9	Get in proper positions for making decisions	Focus on the right area for making decisions
10	Be in the proper angles for decisions	Maintain the proper viewing angle for decisions

more of four dimensions of referee self-efficacy: game knowledge (e.g., *understand the basic strategy of the game*), decision making (e.g., *make critical decisions during competition*), pressure (e.g., *uninfluenced by pressure from players*), and communication (e.g., *communicate effectively with coaches*). Many items are related to more than one dimension: for example, the

item *make critical decisions during competition* relates to all four. Three items were adjusted to improve their relevance to volleyball.

USAV/PAVO First Referee Rating Sheet (See Appendices D, E, & F): The rating sheets were used as the measure of participants' officiating performance. Experienced, high level officials use the rating sheets to evaluate the performance of referees who are candidates for National or Junior National badges. Each candidate is evaluated three times by multiple raters. The officials are rated on seven dimensions of performance: judgment (e.g., consistency), mechanics/signals (e.g., scanning the court before beckoning for the serve), positioning/focus (e.g., watching each ball contact), match control (e.g., warm-up administration), communication with match participants (e.g., demeanor and approachability), communication with officiating team (e.g., interactions with line judges, scorekeeper, and second referee), and professionalism (e.g., appearance). Raters assign points for each dimension of performance (0-15 for each dimension except professionalism, which is 0-10), which is then summed for a total score out of a possible 100 points. Raters also assign a recommended level of play (which ranges from boys' or girls' under-13 to "any level"). The reliability and validity of this instrument have not been investigated in previous research.

Data Collection Procedures

Participants completed a paper survey packet. Permission to use human subjects for this study was obtained from the Institutional Review Board at Michigan State University. I contacted USAV administrators to garner support and cooperation in conducting this research. Participants completed the survey measures at the 2014 USAV Girls' Junior National Championships during a time set aside by tournament administrators. This time was arranged at the beginning of a candidates' informational meeting the evening prior to the participants' first

round of performance evaluations. The lead author was present to conduct the consent process (See Appendix A), administer the measurement instruments and answer any questions the participants may have. Referees were encouraged by USAV to participate, but will not be offered any monetary or material incentive. Participants were given a subject number with their survey packet and instructed to write this number on each of their rating sheets so that their surveys could be matched to their performance evaluations without revealing their identity. The lead researcher captured and entered data from the performance evaluations as the raters turned them in following debriefings with the candidates.

Data Analyses

Data from the paper surveys were entered into excel and double-checked by the researcher and a colleague. Subsequently, all data were loaded into SPSS. Subscale scores for the dimensions and sources of referee efficacy were calculated. Preliminary data analyses, such as multivariate normality, homoscedasticity, univariate normality, outliers, and multicollinearity were conducted as required to screen data before examination (Kline, 1998). Bootstrapping and square root transformations of the data were each attempted after significant skewness and kurtosis were identified for many variables. Although these procedures were successful in producing normality, the results of subsequent analyses did not substantially differ for the adjusted versus the unmodified data.

The data was analyzed in two steps. In the first step, descriptive statistics were calculated for the included variables. Pearson correlations were calculated for all sources and dimensions of officiating efficacy, dimensions of performance, and demographic variables to determine the existence of relationships between key variables. In the second step, multiple regression was used to test the predictive strength of the theorized sources of referee efficacy for dimensions of

referee efficacy, as well as the predictive strength of the sources and dimensions of referee efficacy for the measured dimensions of referee performance.

Originally, it was intended to test the model using structural equation modeling. Due to the lack of significant relationships identified in prior analyses, this final step was abandoned.

CHAPTER 4

RESULTS

The results are presented in three sections. In the first section, the preliminary analyses are presented to evaluate the accuracy and normality of the variables. The second section presents descriptive information of the variables and correlations among the variables. The final section presents the results of testing the individual components of the proposed model using multiple regression.

Preliminary Analyses

The preliminary analyses were conducted to assess the normality and reliability of the variables. To test the assumption of normality of variables, skewness and kurtosis values for each variable were assessed (see Table 2). The assumption of normality can be made if the value of skewness ranges from -1 to +1, and the value of kurtosis ranges from -1 to +2 (Huck, 2004). For the Sport Officials Self-rating Scale, the skewness values of each subscale ranged from -.856 (social support) to .06 (environmental comfort), while kurtosis values ranged from -.892 (environmental comfort) to .72 (vicarious experience). For the Referee Efficacy Scale subscales, skewness values ranged from -1.33 (communication) to -1.02 (pressure), while kurtosis values ranged from .93 (pressure) to 3.35 (communication). For the dimensions of performance on the USAV/PAVO Rating Sheet, skewness values ranged from -1.97 (communication with match officials) to -.28 (positioning/focus), while kurtosis values ranged from -.62 (positioning/focus) to 9.30 (communication with match officials). Reasonable assumptions about normality could be established for each source of referee self-efficacy and for the positioning/focus and communication with match participants dimensions of performance. The assumption of normality could not be met for the remaining dimensions of performance or any of the

dimensions of referee self-efficacy. Attempts to normalize the data using square root transformations and bootstrapping were both successful, but did not substantially alter the results of subsequent statistical tests. Therefore, for ease of interpretation, the statistics reported in this manuscript are based on the original non-transformed, non-bootstrapped data.

Table 2. Summary Statistics for Normality and Internal Consistency

Variables	N	Skewness	Kurtosis	Cronbach's *a*
Judgment	76	-1.04	0.99	0.67
Mechanics	76	-1.33	3.63	0.74
Positioning/Focus	76	-0.28	-0.62	-0.12
Match Control	76	-1.09	2.29	0.84
Communication with match participants	76	-0.73	0.54	0.43
Communication with officials	76	-1.97	9.30	0.29
Professionalism	76	-1.52	4.33	0.54
Overall performance	76	-0.52	0.28	0.61
Social support	76	-0.86	0.50	0.87
Preparation	76	-0.36	-0.12	0.76
Environmental comfort	76	0.06	-0.89	0.83
Situational favorableness	76	-0.22	-0.26	0.71
Past experience	76	-0.80	0.58	0.73
Vicarious experience	76	-0.50	0.72	0.92
Game knowledge	76	-1.27	2.26	0.67
Decision making	76	-1.27	2.62	0.79
Pressure	76	-1.02	0.93	0.82
Communication	76	-1.33	3.35	0.69

For reliability of variables, the Cronbach's alpha values were calculated to evaluate the internal consistency of each variable (see Table 2). Values for the dimensions and sources of referee self-efficacy ranged from .67 (game knowledge dimension of referee self-efficacy) to .92 (vicarious experience as a source of referee self-efficacy). All scale variables met or nearly met Nunnaly's (1978) standard of .72 to be an acceptable reliability coefficient. For the dimensions of performance, taken from the raters' evaluations of the candidates, only match control (.84) and mechanics (.74) met Nunnaly's standard, while several other dimensions showed extremely low reliability. This may be because candidates' performances vary significantly from match to

match within the evaluation period, or because standards are inconsistent across raters.

Descriptive Statistics, Gender Differences, and Differences between Candidate Levels

Descriptive statistics for the dimensions of performance, sources and dimensions of referee self-efficacy, and other potential predictors of performance are presented in Table 3.

Table 3. Descriptive Statistics for Variables

Variables	N	M	SD	Minimum	Maximum
Judgment	76	13.29	0.64	11.20	14.25
Mechanics	76	13.66	0.57	11.20	14.17
Positioning/Focus	76	13.35	0.51	12.00	14.17
Match Control	76	13.27	0.74	10.33	14.60
Communication with match participants	76	14.25	0.38	13.17	15.00
Communication with officials	76	13.84	0.54	11.00	15.00
Professionalism	76	9.54	0.39	7.80	10.00
Overall performance	76	91.30	1.47	86.67	93.83
Social support	76	5.55	0.93	3.20	7.00
Preparation	76	5.77	0.74	3.80	7.00
Environmental comfort	76	3.91	1.47	1.00	7.00
Situational favorableness	76	4.99	0.97	3.00	7.00
Past experience	76	5.47	1.06	2.00	7.00
Vicarious experience	76	4.78	1.16	1.00	7.00
Game knowledge	75	4.50	0.39	3.00	5.00
Decision making	75	4.40	0.49	2.40	5.00
Pressure	73	4.46	0.50	2.80	5.00
Communication	76	4.35	0.43	2.57	5.00
Years officiating	75	14.24	7.35	4.00	35.00
Years coaching	75	13.49	11.35	0.00	38.00
Years playing	75	5.57	7.73	0.00	30.00
Matches officiated last 12 months	75	320.04	219.14	40.00	1500.00
Clinics attended last 12 months	75	4.77	3.50	1.00	20.00

Mean evaluation scores were between 13 and 14 for all dimensions of performance except communication with match participants, which was the highest-scoring dimension at 14.25, and professionalism, which was scored on a 10-point scale rather than 15-point.

Candidates rated social support, preparation, and past experience as the most important sources of confidence in their officiating ability, with mean scores above 5, while environmental comfort was the only source rated below the scale midpoint at 3.9. Mean scores for the four

dimensions of referee self-efficacy were all between 4 and the scale maximum of 5, indicating that the candidates were highly confident in all aspects of their ability.

More variation was observed in the reported volleyball experience of the candidates. Some candidates indicated playing, coaching, and/or officiating volleyball for over 30 years, while others reported as few as 4 years officiating and no experience playing or coaching. Match and training experience over the last year varied widely as well, with a range of 40 to 1500 matches worked and 1 to 20 clinics or training events attended.

One-way ANOVA tests did not reveal any significant differences between genders for any of the variables. However, National-level candidates' performances were rated more highly than Junior National-level candidates' for the dimensions of match control ($F_{(1, 74)} = 9.05$, $p = .004$), communication with match participants ($F_{(1, 74)} = 6.48$, $p = .013$), communication with match officials ($F_{(1, 74)} = 5.14$, $p = .026$), and professionalism ($F_{(1, 74)} = 11.12$, $p = .001$), as well as overall ($F_{(1, 74)} = 11.61$, $p = .001$). Compared to Junior National-level candidates, National-level candidates also reported more years playing volleyball ($F_{(1, 73)} = 7.63$, $p = .007$), but fewer matches officiated in the past 12 months ($F_{(1, 74)} = 4.98$, $p = .029$).

Correlations among the Variables

Pearson correlations were calculated in order to find which sources of referee self-efficacy were related to dimensions of self-efficacy (see Table 4), and which potential predictors of referee performance were related to the various dimensions of performance upon which the candidates were evaluated (see Table 5).

Table 4 displays the correlations of the sources of referee self-efficacy with the four dimensions of referee self-efficacy. Only one significant relationship emerged: preparation was

significantly, positively related to game knowledge, $r = .24$, $p = .039$. There may also be a modest relationship between preparation and pressure, but it was not statistically significant, $r = .23$, $p = .055$. Individuals reporting high self-efficacy in the game knowledge dimension reported drawing confidence from their preparation. No other significant relationships or strong trends were observed.

Table 4. Pearson Correlation Coefficients between Sources and Dimensions of Referee Self-Efficacy

Sources of Efficacy	Game Knowledge	Decision-Making	Pressure	Communication
Social Support	-.02	-.09	-.06	-.01
Preparation	.24*	.19	.23[1]	.18
Environmental Comfort	-.06	-.04	-.04	-.00
Situational Favorableness	-.02	-.07	-.05	.04
Past Experience	.09	.11	.15	.15
Vicarious Experience	.14	-.02	.08	.12
Years Officiating	.12	.14	.11	.15
Years Playing	.13	.20	.17	.14
Years Coaching	.16	.17	.15	.15
Matches last 12 months	.10	.03	-.07	-.03
Clinics last 12 months	.00	.08	.05	.01

* $p < .05$, [1] $p = .055$ (2-tailed)

Table 5. Pearson Correlation Coefficients between Performance Predictors and Performance Ratings

Predictors	Ovr	Judg	Mec	Pos	MC	ComP	ComO	Prof
Social support	-.18	.02	-.03	.15	-.39**	-.14	-.06	-.01
Preparation	-.17	.06	.00	-.06	-.29*	.01	-.11	-.13
Environmental comfort	-.21[1]	-.13	-.00	-.12	-.27*	-.02	-.10	.02
Situational favorableness	-.02	.06	-.08	.08	-.23[1]	-.11	.07	.12
Past experience	-.15	-.01	-.03	-.07	-.10	-.12	-.11	-.18
Vicarious experience	-.16	-.03	-.06	-.09	-.15	-.03	-.19	.04
Game knowledge	.19	-.11	.02	.14	-.04	.20	.13	-.04
Decision making	.20[1]	-.07	.03	.09	.04	.19	.15	-.05
Pressure	.12	-.05	.04	.10	.01	.10	.18	-.11
Communication	.12	-.09	.04	.10	-.04	.12	.11	-.03
Years officiating	-.11	-.10	-.17	.00	-.07	.10	.11	.10
Years playing	.08	-.07	-.15	-.03	.06	.19	-.07	.02
Years coaching	.02	-.08	-.21[1]	.01	-.05	.26*	-.20[1]	-.10
Matches last 12 months	-.05	.00	-.05	-.00	.03	-.06	-.07	.01
Clinics last 12 months	.12	-.10	.13	-.02	.03	.21[1]	.00	.24*

** $p < .01$, * $p < .05$, [1] $p < .10$ (2-tailed)

Note: Ovr = Overall, Judg = Judgment, Mec = Mechanics/Signals, Pos = Positioning/Focus, MC = Match Control, ComP = Communication with Match Participants, ComO = Communication with Officiating Team, Prof = Professionalism

Table 5 displays the correlations of the sources of referee self-efficacy, dimensions of referee self-efficacy, and other potential performance predictors with the various dimensions of the candidates' performance evaluation ratings. There were no statistically significant relationships between any predictors and the overall performance score, but two predictors showed statistically insignificant trends: environmental comfort as a source of self-efficacy trended negatively with overall performance score, $r = -.21$, $p = .066$, while the decision-making efficacy dimension trended positively, $r = .20$, $p = .094$. Referees who drew more confidence from environmental comfort may have performed worse overall, while candidates who reported high levels of self-efficacy for decision-making may have performed better.

The match control dimension of performance was significantly, negatively correlated with three sources of referee efficacy: social support, $r = -.38$, $p < .001$, preparation, $r = -.28$, $p = .013$, and environmental comfort, $r = -.27$, $p = .020$. There was also a negative trend between match control and situational favorableness, $r = -.23$, $p = .051$. Candidates who reported drawing more confidence from these sources tended to score lower on the match control dimension of performance. Years of experience coaching volleyball was significantly, positively correlated to performance in the communication with match participants dimension, $r = .26$, $p = .024$. There were also negative trends for experience coaching with the mechanics dimension of performance, $r = -.21$, $p = .076$, and communication with other officials, $r = -.20$, $p = .083$. Referees with more coaching experience performed better when working with players and coaches, but may have showed weaker officiating mechanics and ability to communicate with their fellow officials. Clinics attended in the past 12 months was significantly, positively related to professional performance, $r = .24$, $p = .043$, and showed a positive trend for communication with match participants, $r = .21$, $p = .076$. Candidates who attended more clinics and training

events were graded higher on their professionalism during matches, and may have showed better ability to communicate with players and coaches. No other relationships or strong trends were observed between sources and dimensions of referee self-efficacy, experience and training, and performance ratings.

Tests of Components of the Proposed Model

The first stage of the model was tested by four univariate multiple regressions using years of experience, matches and clinics/training events in the past year, and the six sources of referee self-efficacy as predictors and the four dimensions of referee self-efficacy as criterion variables. The results of the multiple regression analysis for each REFS subscale are summarized in Table 6.

Table 6. The Predictability of Referee Self-Efficacy Sources and Experience on Referee Self-Efficacy Dimensions

Predictors	Game Knowledge	Decision-Making	Pressure	Communication
Social Support (β)	-.12	-.17	-.16	-.15
Preparation (β)	.27[1]	.29*	.38*	.18
Environmental Comfort (β)	-.21	-.15	-.20	-.21
Situational Favorableness (β)	-.06	-.10	-.05	.05
Past Experience (β)	.18	.30[1]	.34*	.20
Vicarious Experience (β)	.07	-.15	-.29[1]	.02
Years Officiating (β)	.17	.17	.14	.15
Years Playing (β)	.04	.10	.06	.08
Years Coaching (β)	.09	.06	.06	.09
Matches last 12 months (β)	.11	.05	-.06	-.04
Clinics last 12 months (β)	-.02	.09	.11	.03

* p <.05, [1] p < .10 (2-tailed)

The predictors in the regression models did not significantly predict referee self-efficacy for the dimensions of game knowledge ($F_{(11, 62)} = 1.06$, $p = .406$), decision-making ($F_{(11, 62)} = 1.23$, $p = .285$), pressure ($F_{(11, 60)} = 1.57$, $p = .130$), or communication ($F_{(11, 63)} = .74$, $p = .698$). With regard to individual predictors, preparation significantly predicted decision-making ($\beta = .29$, $p = .046$) and pressure ($\beta = .38$, $p = .009$), and may predict game knowledge ($\beta = .27$, $p = .059$). Past experience predicted pressure ($\beta = .34$, $p = .039$) and may predict decision-making (β

= .30, p = .058), and vicarious experience may predict pressure (β = -.29, p = .053). No other significant or near-significant beta weights were observed. A simplified regression model including only the self-efficacy sources preparation, past experience, and vicarious experience did significantly predict the self-efficacy dimension pressure, ($F_{(3, 69)}$ = 3.60, p = .018), accounting for 9.8% of variance based on adjusted R^2. Thus, the first stage of the proposed model was not well-supported by the data. While one dimension of referee self-efficacy may be weakly associated with a handful of predictors, by and large the proposed predictors (experience and sources of referee self-efficacy) were not observed to relate to the four dimensions of referee self-efficacy.

The second part of the model was tested by eight univariate multiple regressions using the dimensions of referee self-efficacy and the experience variables (years of experience, matches in the last 12 months, and clinics in the last 12 months) as predictors and the performance dimensions and overall performance score as criterion variables. The results are displayed in Table 7.

Table 7. The Predictability of Referee Self-Efficacy Dimensions and Experience on Performance

Predictors	Ovr	Judg	Mec	Pos	MC	ComP	ComO	Prof
Game knowledge	.07	-.21	-.10	.18	-.33	.10	.13	-.06
Decision making	.62	.12	.12	-.22	.56	.59	-.04	.29
Pressure	-.40	-.04	-.04	.21	-.32	-.50[1]	.24	-.36
Communication	-.08	.06	.13	.01	.08	-.06	-.12	.03
Years officiating	-.17	-.11	-.14	.01	-.12	.01	.13	.14
Years playing	.05	-.01	-.09	-.11	.08	.08	-.08	.00
Years coaching	-.04	-.04	-.20	.04	-.08	.17	-.22	-.14
Matches last 12 months	-.16	.00	-.06	.03	-.06	-.22[1]	-.02	-.02
Clinics last 12 months	.11	-.11	.15	.00	.01	.20	.05	.27*

* p < .05 , [1] p < .10 (2-tailed)

Note: Ovr = Overall, Judg = Judgment, Mec = Mechanics/Signals, Pos = Positioning/Focus, MC = Match Control, ComP = Communication with Match Participants, ComO = Communication with Officiating Team, Prof = Professionalism

The predictors in the regression model did not significantly predict referee performance overall ($F_{(9, 62)} = .99$, $p = .458$) or for the dimensions judgment ($F_{(9, 62)} = .25$, $p = .984$), mechanics ($F_{(9, 62)} = .79$, $p = .627$), positioning/focus ($F_{(9, 62)} = .28$, $p = .978$), match control ($F_{(9, 62)} = .38$, $p = .937$) communication with match participants ($F_{(9, 62)} = 1.85$, $p = .077$), communication with other officials ($F_{(9, 62)} = .82$, $p = .600$), or professionalism ($F_{(9, 62)} = .90$, $p = .531$). Because communication with match participants was near significance, a simplified regression model including only the game knowledge, decision-making, and pressure dimensions of referee self-efficacy, with communication with match participants as the criterion variable, was attempted. This trimmed model was also insignificant, ($F_{(3, 69)} = 1.72$, $p = .171$). Only three predictors had standardized coefficients for individual dimensions of performance that were significant or approached significance: pressure, for communication with match participants ($\beta = -.50$, $p = .090$); matches in the last 12 months, also for communication with match participants ($\beta = -.22$, $p = .074$); and clinics/training events attended in the last twelve months, for professionalism ($\beta = .27$, $p = .039$).

Due to the lack of functional connections between and within stages of the model, no further modeling techniques or tests were conducted.

CHAPTER 5

DISCUSSION

Despite the importance of sport officials' performance to the experience of match participants and other actors in competitive contexts, little is known about the root causes of that performance. Past research has identified experience as a primary predictor of performance in referees, but until recently no psychological factors had been scientifically associated with performance. Guillen and Feltz (2011) suggested that self-efficacy might be a psychological construct of interest for referees, and Myers et al. (2012) provided an instrument for measuring self-efficacy specifically in sport officials. This research is the first attempt at evaluating the relationship between self-efficacy and performance in sport officials.

The main purpose of this study was to examine the predictive strength of referee self-efficacy on officiating performance in volleyball referees. None of the dimensions of performance were significantly predicted by regression models including the four dimensions of referee self-efficacy (game knowledge, decision-making, pressure, and communication) and the experience variables collected in the study (years of experience officiating, coaching, and playing, matches officiated in the last 12 months, and clinics or training events attended in the last 12 months). There was a significant relationship between clinics attended and the professionalism dimension of performance. This is a rather curious result, as the impression given by the rating team was that the professionalism score was something of a "throwaway," used to get the overall score to a point where the rater was happy after they could not find points to deduct in other categories. One explanation may be that officials who are more professional may invest more time and money into attending, and this professional approach may come out in their demeanor and performance in matches. As raters often scored candidates after debriefing

them and providing feedback, the professionalism score may also have been influenced by the candidates' reactions to the raters' evaluations. Finally, this result may be the result of bias on the raters' part toward individuals who are more actively involved in the development process and officials' community, such that they rate more highly individuals who they are more familiar with, or seek to reward candidates who may have shown more investment in their own development. However, the rating team generally seemed to avoid assigning raters to evaluate individuals from their own home region, so this explanation may not apply in most cases.

There were reasonably strong trends of association between the pressure dimension of self-efficacy and matches officiated, and the communication with match participants dimension of performance. Communication skills developing with experience is no surprise, but at first glance the relationship between pressure and communication might seem unintuitive: why would referees' self-efficacy for pressure, rather than communication, influence their ability to communicate with coaches and players, and why is that relationship negative? The USAV rating sheet lists the following items of evaluation for communication with match participants:

1. Respectful, dignified manner
2. Demeanor, approachability
3. Communication with team members
4. Acknowledgement of coaches

It may be that players and coaches are the primary source of pressure on referees in this environment, and the official's ability to maintain their manner and demeanor under that pressure is more relevant to the rater's evaluation in this dimension than their ability to communicate outward to the match participants. The negative direction of the relationship may reflect a

misalignment of the candidates' impressions of their ability to handle pressure with the realities of their behavior; perhaps some referees who think they are skilled in handling pressure do so by "brushing it off" or ignoring match participants' attempts at communication, rather than displaying the attributes and behaviors for which the raters are instructed to look.

A secondary purpose of this research was to investigate the connections between the sources and the dimensions of referee self-efficacy. Two sources significantly predicted one or more dimensions: preparation was related to decision-making and pressure efficacy, while past experience was only significantly related to pressure efficacy. Predictive relationships may also exist between preparation and game knowledge efficacy, past experience and decision-making efficacy, and vicarious experience and pressure efficacy, but these were not statistically significant. It should not be surprising that, in data light on strong associations, preparation and past experience emerged as the only notable predictors: in developing the Sources of Sport Confidence Questionnaire, Vealey et al. (1998) identified preparation as the most salient source of self-confidence in athletes, and preparation and experience are consistently forwarded as two of the most important efficacy sources (e.g. Bandura, 1977, 1986). In the Myers et al. (2012) study, preparation predicted all four dimensions of referee efficacy, environmental comfort and situational favorableness predicted decision-making, pressure, and communication efficacy, and past accomplishments and vicarious experience predicted game knowledge, decision-making, and pressure efficacy. The results were similar in that social support was not found to significantly predict any dimension. Myers and colleagues found that years of experience and highest level of experience also predicted all four dimensions of referee self-efficacy, but there was no such relationship for the experience measures in the current study.

One explanation for the lack of significant results from this study is the absence of diversity in the sample. The candidates were of similar skill levels—hence their congruent candidacies—and the vast majority of them passed their evaluations. My impression was that the regions actively winnow out local referees who they do not feel are ready to move up yet, forwarding to the evaluation process only those whom they believe to be skilled and experienced enough to advance. The spread of performance evaluation scores was relatively narrow, especially when comparing candidates' average scores across multiple evaluations. Indeed, overall performance scores low enough to fail the evaluation were by some measures statistical outliers. Additionally, the candidates as a group were very confident in their abilities, perhaps as a result of the aforementioned winnowing process. Scores on the REFS dimensions averaged well over 4 on a 1-5 scale, with standard deviations of less than 0.5. Including a more diverse range of ability levels might have generated more significant results. Efficacy scores in the Myers et al. (2012) study were similarly high, but that study made no attempt to use them to discriminate between outcomes like performance.

Problems with the outcome measure might also explain the lack of significant findings. The USAV rating sheets were used as the outcome measure in this study because they provided an opportunity to collect a standardized performance score for a relatively large number of referees. Introduction of a more scientifically validated measure would have been optimal, but impractical. I did not anticipate the degree to which the evaluation process would be compromised. While all raters used the rubric and made comments relating to individual dimensions of performance, some did not assign category scores, choosing instead to simply assign an overall score out of 100. As mentioned before, some that did give category scores used

professionalism as something of a dummy score, docking a point if they thought the overall score too high after summing the other dimensions.

More concerning, the evaluation process was rife with examples of "poisoning the well." Raters would frequently discuss their observations with the next rater scheduled to evaluate a candidate. While this might have made the raters' jobs easier, it must call into question the value of the rating sheet as a valid metric for performance. This was symptomatic of what I view as a larger problem: the evaluation process was intertwined with training of the candidates in ways which damaged the usefulness of the evaluation. Candidates were debriefed by their raters following each two-match round of evaluations, and then in subsequent matches judged—at least in part—on their efforts to address issues brought up during the debriefing, rather than their performance in that match alone. This phenomenon was most evident in nightly meetings of the rating team, in which the evaluators would discuss each candidate one by one, noting their current standing with regard to passing or failing, and warning fellow raters what to look out for the next day. Because opportunities for training of the candidates are limited, it may be beneficial or even necessary from USAV's perspective for the evaluation process to involve some training. From a researcher's perspective, however, this mixture is a problem because the outcome variables, which were intended to represent objective performance, instead to some degree represent the candidates' response to feedback.

Finally, the context in which the referees were evaluated must be considered. The tournament was staged in a convention center with several dozen courts in a single exhibition hall. The event space was cramped, with parents and spectators very close to the playing area, and with the number of whistles and shouting athletes it was extremely loud. Referees worked long days in this environment, and evening social events and unfamiliar sleeping arrangements

may have limited their quality rest. This was a stressful environment for the candidates, compounded by their knowledge that they were being evaluated for professional advancement. With that in mind, it makes sense that the referees' self-efficacy for coping with pressure was one of the more salient (if statistically insignificant) predictors of some performance dimensions. Alternatively, one could argue that this may be the "natural" context for volleyball referees at this level—while referees often work high school or college contests, they also frequently call matches at large-scale events like this one. In that case, perhaps the evaluation context was appropriate to the demands of the avocation. The working situation for the evaluation team should also be considered: the raters worked for even longer hours than the candidates, and it is unclear whether their scoring may have changed over the course of several days of evaluations. The stress and distractions of the event space may impact the raters as well, and all of the scores are confounded by variation in level of play and difficulty of officiating unique events from match to match.

Limitations and Future Directions

The primary limitation of this study was its small number of participants. Myers et al. (2012) suggested N=250 to use most parameters of the REFS, and N=400 to use all parameters. Seventy-six referees participated in the study, representing nearly all of the candidates for the year. A similar study conducted with officials from another sport might incorporate a larger sample size to improve power. The current study's reliance on a single, standardized outcome measure for performance might have been a strength; unfortunately, the compromised objectivity of that measure may mean that the lack of other measures to interpret is instead a limitation. Replication with a larger group might necessitate incorporation of multiple performance measures or, more ambitiously, development of a more broadly applicable instrument. Another

limitation of this study was the homogeneity of participants with regard to skill and confidence levels. Future research might involve officials at a wider variety of skill and experience levels.

As this project was observational in nature, causal inference regarding the relationship between self-efficacy and performance cannot be drawn from the results. Future experimental research could use bogus feedback to manipulate officials' self-efficacy in a laboratory-reproducible task, such as calling balls and strikes, or making offside decisions based on video from a soccer match. Tasks of this nature have been shown to discriminate between referee skill levels or specializations in the past, but never used in conjunction with manipulation of psychological constructs like self-efficacy.

Conclusion

No relationship was found in this thesis between self-efficacy and performance in high-level volleyball officials. There also was little relation found between referees' level of experience and performance in their evaluations, and few connections between previously identified sources of referee confidence and dimensions of referee self-efficacy. These null findings may be due to lack of variance in ability and confidence on the part of the referees, or produced by an evaluation system which is designed to teach candidates, and pass most of them, rather than explicitly evaluate their performance. The results of this study do replicate those of the Myers et al. (2012) study in identifying preparation and past experience as the most important sources of referee efficacy. Future research might incorporate a more objective evaluation system, a wider variety of referee skill and experience levels, officials from different sports, and/or experimental manipulation of self-efficacy during simulated officiating tasks in a laboratory setting.

APPENDICES

Appendix A: Consent Form

Volleyball Officials Research Study Participant Consent Form

You are being asked to participate in a research study conducted by Masters' degree student Ben Spencer under the supervision of Deborah Feltz, Ph.D. from Michigan State University. This study is to investigate the contribution of self-confidence to match performance in volleyball officials. You have been identified as a potential participant in this study because you are a volleyball referee. You must be at least 18 years old to participate in this research.

Procedure: As part of this research, you will be asked to complete a short packet of surveys designed to provide background information about yourself as an official, your self-confidence for the officiating task, and your sources of self-confidence. It should take approximately 10 minutes to complete. Additionally, copies of your completed rating sheets will be taken as a performance measure.

Benefits: We believe that the study results will have practical applications for volleyball officials to improve their performance. Additionally, the information gained from this study will increase our understanding of different aspects of sports officiating.

Risks: There are no known physical, legal, or economic risks associated with this study. None of the questions address sensitive issues regarding personal beliefs, behaviors, experiences or attitudes.

Voluntary Participation: Participation in this research project is completely voluntary. You have the right to say no. You may change your mind at any time and withdraw. You may choose not to answer specific questions or to stop participating at any time. Whether you choose to participate or not will have no affect on your evaluation.

Confidentiality: Your participation in this study will remain confidential. The principal investigator, secondary investigator, and the IRB will have access to the research data. It will be kept in a locked file cabinet and on a password protected computer. All collected data will be de-identified and analyzed at the group level to ensure the confidentiality of individual responses. Your confidentiality will be protected to the maximum extent allowable by law.

Contact and Questions: If you have concerns or questions about this study, such as scientific issues, how to do any part of it, or to report an injury, please contact the researchers: Ben Spencer (spenc291@msu.edu), (402) 429-1500; or Deborah L. Feltz, Ph.D. (dfeltz@msu.edu), (517) 355-4732, or by regular mail at: Michigan State University, 134 IM Circle, East Lansing, MI 48824.

If you have questions or concerns about your role and rights as a research participant, would like to obtain information or offer input, or would like to register a complaint about this study, you may contact, anonymously if you wish, the Michigan State University's Human Research Protection Program at 517-355-2180, Fax 517-432-4503, or e-mail irb@msu.edu or regular mail at 207 Olds Hall, MSU, East Lansing, MI 48824.

Statement of Assent/Consent: Your signature below means that you voluntarily agree to participate in this research study and release copies of your rating sheets to be used in this research.

_____ _____

Signature Date

Appendix B: Demographics and Sport Official Self-Rating Scale

Sport Officials Self-Rating Scale

Think back to times when you felt very confident when officiating in your sport. What things made you feel confident? What things helped you believe in your abilities and gave you confidence that you would perform successfully?

Participant ID #: _____ Race: White/Caucasian □ Black/African American □

Age: Gender: Female □ Male□ Asian/Pacific Islander □ Hispanic □

USAV Region: _____ American Indian/Alaskan Native □ Other □

Is volleyball the primary sport you officiate?: Yes □ No □

Do you referee other sports?: Yes □ No □ List: _____

Years experience officiating volleyball: ___ playing volleyball: ___ coaching volleyball: ___

Highest level as referee: ___ as player: ___ as coach: ___

Approximately how many matches have you officiated in the past 12 months? ___

In how many clinics or training events have you participated in the last 12 months? ___

Listed below are some things that may help officials feel confident in performance situations. For each statement, check the number that indicates HOW IMPORTANT THAT IS IN HELPING YOU FEEL CONFIDENT IN OFFICIATING. Please respond to every question even though they may seem repetitive. There are no right or wrong answers because every official is different. Please be honest—your answers will be kept completely confidential.

I gain confidence in officiating when I...

	Not at all important 1		Not very important 2		Slightly important 3		Of average importance 4		Very important 5		Extremely important 6		Of highest importance 7	
1.	Get positive feedback from other officials	1	2	3	4	5	6	7						
2.	Keep my focus on the game	1	2	3	4	5	6	7						
3.	Officiate in a venue I like	1	2	3	4	5	6	7						
4.	Am familiar with officials I will officiate with	1	2	3	4	5	6	7						
5.	Performed well in previous contests	1	2	3	4	5	6	7						
6.	Know I have support from other officials in my sport	1	2	3	4	5	6	7						
7.	See successful officiating by other officials in my sport	1	2	3	4	5	6	7						
8.	Know I'm mentally prepared for the contest	1	2	3	4	5	6	7						
9.	Watch another official I admire perform successfully	1	2	3	4	5	6	7						
10.	Am assigned a match/game I feel qualified for	1	2	3	4	5	6	7						
11.	Feel comfortable in the venue in which I'm officiating	1	2	3	4	5	6	7						
12.	Made good decisions in previous contests	1	2	3	4	5	6	7						

13. Am encouraged by other officials	1	2	3	4	5	6	7
14. Watch another official perform well	1	2	3	4	5	6	7
15. Venue conditions are favourable	1	2	3	4	5	6	7
16. Prepare myself physically and mentally for a contest	1	2	3	4	5	6	7
17. Like the venue where I am officiating	1	2	3	4	5	6	7
18. Have performed well in difficult contests	1	2	3	4	5	6	7
19. Get positive feedback from evaluators of my officiating	1	2	3	4	5	6	7
20. Watch well-officiated contests	1	2	3	4	5	6	7
21. Believe in my ability to give maximum concentration in a contest	1	2	3	4	5	6	7
22. Receive support and encouragement from other officials	1	2	3	4	5	6	7
23. Watch officials who are at my level perform well	1	2	3	4	5	6	7
24. Am assigned to officiate with a qualified partner	1	2	3	4	5	6	7
25. Am in good physical condition	1	2	3	4	5	6	7

Appendix C: Refficacy Questionnaire

REFFICACY QUESTIONNAIRE
Referee confidence refers to the extent to which referees believe that they have the capacity to perform successfully in their job. **Think about how self-confident you are when you officiating. Truthfully respond to the questions below based on how confident you feel about officiating. There are no correct answers. Please be honest—your answers will be kept completely confidential. Circle the number which corresponds to your feelings of self-confidence.**

In the context of performing your referee job, how confident are you in your ability to	**Low**		**Medium**		**High**
1. Understand the rules of your sport	1	2	3	4	5
2. Know when and how to call more or fewer faults/infractions to control the flow of the game	1	2	3	4	5
3. Demonstrate poise under pressure	1	2	3	4	5
4. Communicate effectively with coaches	1	2	3	4	5
5. Stay up with the play	1	2	3	4	5
6. Think and respond successfully during competition	1	2	3	4	5
7. Resolve disputes	1	2	3	4	5
8. Apply the rules accurately	1	2	3	4	5
9. Focus on the right area for making decisions	1	2	3	4	5
10. Maintain the proper viewing angle for decisions	1	2	3	4	5
11. Make critical decisions during match (game/competition)	1	2	3	4	5
12. Be in control of the game	1	2	3	4	5
13. Be successful as a referee at your current level	1	2	3	4	5
14. Concentrate well enough to be successful	1	2	3	4	5
15. Communicate effectively with partners	1	2	3	4	5
16. Recover from making a bad call to make a “correct” call	1	2	3	4	5
17. Consistently be successful in making correct decisions	1	2	3	4	5
18. Uninfluenced by pressure from players	1	2	3	4	5
19. Handle unexpected situations	1	2	3	4	5
20. Demonstrate effective teamwork with partners	1	2	3	4	5
21. Recognize your own mistakes	1	2	3	4	5
22. Uninfluenced by pressure from spectators	1	2	3	4	5
23. Adapt to different game situations and still be successful	1	2	3	4	5
24. Achieve your professional goals as a referee	1	2	3	4	5
25. Know and understand the basic strategy of the game	1	2	3	4	5
26. Communicate effectively with players	1	2	3	4	5
27. Be in good physical condition	1	2	3	4	5
28. Handle challenges about decisions appropriately	1	2	3	4	5
29. Demonstrate decisiveness	1	2	3	4	5
30. Anticipate game situations	1	2	3	4	5
31. Communicate effectively with auxiliary game personnel (e.g., video reviewer, scorekeepers, timekeepers, goal judges, etc)	1	2	3	4	5

32. Be firm in your decisions	1	2	3	4	5
33. Know and understand proper officiating mechanics	1	2	3	4	5
34. Know all the rules of your sport	1	2	3	4	5
35. Make quick decisions	1	2	3	4	5
36. Not let a bad call affect your next call	1	2	3	4	5
37. Demonstrate accurate judgement	1	2	3	4	5
38. Be successful even when the crowd is against you	1	2	3	4	5
39. Uninfluenced by pressure from coaches	1	2	3	4	5

Appendix D: USAV/PAVO Referee Rating Sheets

USAV/PAVO FIRST REFEREE RATING SHEET

Date: _____ **Candidate:** _____ **Total Score:** _____

Event/Site: _____ **Partner:** _____

Level of Play: _____ **Time/Court:** _____ **Teams:** _____

Rater: _____ **Match Scores:** _____

Match Difficulty: _____ 1 - Few decisions; little pressure 2 - Average match; 3 - Decisions numerous and difficult; high pressure

Key to Match Situations

A	Attack	IYC	Individual Yellow Card	SR	Serve Receive	**Rater's Judgment**	
BRA	Illegal Attack	N	Net	S	Set (throw)	+	Good call
BRB	Illegal Block	OP	Out of Position	S2	Set (double hit)	—	Missed call
B	Block	P	Pass (held ball)	DP	Delay Penalty	X	Should not have been called
CL	Center Line	P2	Pass (double hit)	DW	Delay Warning		
IR	Improper Request	RON	Reaching Over the Net	T	Tip	0	Appropriate no call
IRC	Individual Red Card	SC	Screening				

15 pts.	**A. Judgment** 1. Consistency: a) Set to set b) Team to team c) Skill level to skill level d) Position to position e) Action to action 2. Setting 3. Passing 4. First team contacts 5. Tips and attacks 6. Net/block recoveries 7. Blocking 8. Unorthodox techniques 9. Illegal attacks or blocks 10. Net faults/non-interfering contact 11. Antenna decisions 12. Touches 13. Screening 14. Position/rotation faults 15. Adjustment to level of play 16. Ability to stay with the play 17. Decisiveness 18. Anticipation	
15 pts.	**B. Mechanics/Signals** 1. Scan before beckon 2. Signals 3. Techniques 4. Signal sequence 5. Whistle quality and technique 6. Reaction time 7. Mimics appropriate signals initiated by the second referee	

Comments/Strengths:

1.

2.

3.

LEVEL OF PLAY RECOMMENDED FOR THIS REFEREE:

_____

Candidate's Signature: _____ **Date:**_____

Revised September 2013

Figure 3: USAV/PAVO First Referee Rating Sheet

Figure 3 (cont'd)

USAV/PAVO FIRST REFEREE RATING SHEET (cont'd.)

Key to Match Situations

A	Attack	IYC	Individual Yellow Card	SR	Serve Receive	+	Good call
BRA	Illegal Attack	N	Net	S	Set (throw)	—	Missed call
BRB	Illegal Block	OP	Out of Position	S2	Set (double hit)	X	Should not have been called
B	Block	P	Pass (held ball)	DP	Delay Penalty		
CL	Center Line	P2	Pass (double hit)	DW	Delay Warning		
IR	Improper Request	RON	Reaching Over the Net	T	Tip	0	Appropriate no call
IRC	Individual Red Card	SC	Screening				

Rater's Judgment

15 pts.	**C. Positioning/Focus** 1. Changes focal point appropriately 2. Watches each ball contact 3. Eye movement 4. Quick adjustments 5. Position	
15 pts.	**D. Match Control** 1. Site inspection 2. Warm-up administration 3. Match protocol, time management 4. Verbal warnings 5. Sanctions 6. Protects officiating team 7. Bench awareness and control 8. Court awareness 9. Assures participant safety 10. Match tempo	
15 pts.	**E. Communication with Match Participants** 1. Pre-match meeting & coin toss 2. Respectful, dignified manner 3. Demeanor, approachability 4. Communication with team members 5. Acknowledgement of coaches	
15 pts.	**F. Communication with Officiating Team** 1. Pre-match briefings 2. Communication skills, engages crew 3. Eye contact; centering 4. Interactions with – • Line judges • Scorekeeper • Second referee	
10 pts.	**G. Professionalism** 1. Appearance 2. Comportment 3. Recovery/Unusual situations 4. Presence 5. Confidence 6. Cooperation 7. Attitude	

Comments/Areas for Improvement/Points of Emphasis:

1.

2.

3.

Revised September 2013

USAV/PAVO SECOND REFEREE RATING SHEET

Date: _____ **Candidate:** _____ **Total Score:** _____

Event/Site: _____ **Partner:** _____

Level of Play: _____ **Time/Court:** _____ **Teams:** _____

Rater: _____ **Match Scores:** _____

Match Difficulty: _____ 1 - Few decisions; little pressure 2 - Average match; 3 - Decisions numerous and difficult; high pressure

Key to Match Situations

A	Attack	IYC	Individual Yellow Card	SR	Serve Receive	**Rater's Judgment**	
BRA	Illegal Attack	N	Net	S	Set (throw)	+	Good call
BRB	Illegal Block	OP	Out of Position	S2	Set (double hit)	—	Missed call
B	Block	P	Pass (held ball)	DP	Delay Penalty	X	Should not have been called
CL	Center Line	P2	Pass (double hit)	DW	Delay Warning		
IR	Improper Request	RON	Reaching Over the Net	T	Tip	0	Appropriate no call
IRC	Individual Red Card	SC	Screening				

15 pts.	**A. Judgment** 1. Net faults/non-interfering contacts 2. Center line faults 3. Position faults on receiving team 4. Illegal attacks or blocks 5. Touches 6. Antenna decisions 7. Anticipation 8. Adjusts to level of competition	
15 pts.	**B. Mechanics/Signals** 1. Signals 2. Techniques 3. Signal sequence 4. Mimics appropriate signals initiated by the first referee 5. Whistle quality and technique 6. Reaction time 7. Court scan	
15 pts.	**C. Positioning/Focus** 1. Positioned to see receiving team 2. Positioned on blocking side during play 3. Proper distance from net support 4. Timely & quick transitions 5. Secondary transitions 6. Clear view of blockers and attackers 7. Ability to see play develop 8. Clearly visible to first referee when needed 9. Eye movement	

Comments/Strengths:

1.

2.

3.

LEVEL OF PLAY RECOMMENDED FOR THIS REFEREE:

_____

Candidate's Signature: _____ **Date:** _____

Revised September 2013

Figure 4: USAV/PAVO Second Referee Rating Sheet

Figure 4 (cont'd)

USAV/PAVO SECOND REFEREE RATING SHEET (cont'd.)

Key to Match Situations

A	Attack	IYC	Individual Yellow Card	SR	Serve Receive	+	Good call
BRA	Illegal Attack	N	Net	S	Set (throw)	—	Missed call
BRB	Illegal Block	OP	Out of Position	S2	Set (double hit)	X	Should not have been called
B	Block	P	Pass (held ball)	DP	Delay Penalty		
CL	Center Line	P2	Pass (double hit)	DW	Delay Warning		
IR	Improper Request	RON	Reaching Over the Net	T	Tip	0	Appropriate no call
IRC	Individual Red Card	SC	Screening				

Rater's Judgment

15 pts.	**D. Match Control** 1. Match protocol, time management 2. Warm-up administration 3. Pre-game lineup check 4. Decisiveness 5. Facilitator 6. Court awareness 7. Responsive to bench needs 8. Substitution procedures 9. Timeout procedures 10. Bench control 11. Verbal warnings 12. Sanctions 13. Assures participant safety 14. Protects officiating team 15. Match tempo	
15 pts.	**E. Communication with Match Participants** 1. Respectful, dignified manner 2. Demeanor, approachability 3. Communication with team members 4. Acknowledgement of coaches	
15 pts.	**F. Communication with Officiating Team** 1. Pre-match briefing with first referee 2. Pre-match briefing with scorekeeper and other staff at score table 3. Communication skills during the match; engages crew 4. Eye contact; centering 5. Discreet assistance to first referee 6. Supportive, encouraging 7. Awareness of line judges 8. Supervision of auxiliary personnel (speed wipers/ball crew)	
10 pts.	**G. Professionalism** 1. Appearance 2. Comportment 3. Recovery/Unusual situations 4. Poise 5. Confidence 6. Cooperation 7. Attitude	

Comments/Areas for Improvement/Points of Emphasis:

1.

2.

3.

Revised September 2013

Appendix E: Practical Rating Sheet Instructions

PRACTICAL RATING SHEET INSTRUCTIONS

Forms Used During a Practical Rating Session

- **First and Second Referee Criteria Sheets -** These forms detail the various elements that rates should be looking for when evaluating a candidate. The detailed criteria correlate with the abbreviated phrases on the actual rating sheet. Each rater should be very familiar with the criteria listed, and each candidate should be given a copy prior to the evaluation session. It is informational only – no comments or notes need be written on this form by the rater.
- **First and Second Referee Rating Sheets -** The rating sheets themselves are separated into several major categories, with an abbreviated version of the criteria beneath each heading. The information at the top of the form should be filled out completely. Comments associated with the category should be written in the boxes on the right hand side, and are needed any time points are deducted from a category. Raters should include positive observations as well as constructive comments. Suggested abbreviations for common calls and situations are given in the key at the top of each page.
- **List of Critical Errors -** The "Standards for Officials' Ratings" in the PAVO Board Handbook explains the minimum practical evaluation score that is necessary for each rating level. However, there are some errors that are so egregious that, if they occur and are confirmed, the candidate should not be considered for an advanced rating at that time. This list should be shared with raters and candidates prior to the evaluation session.

Rating Sheet Example

Please review the completed example sheets enclosed. As you will notice, the rater used a combination of abbreviations and play-by-play comments to record the officiating in the match as it occurred. On the first referee rating sheet, the "Judgment" section is the most difficult. The rater should record **decisions**, not just calls they disagree with. A **decision** can have four possible outcomes in terms of this evaluation. The rater may judge that:

- a call made was appropriate (commonly noted with a "+" sign)
- a call was not made that should have been, or a fault was missed (commonly noted with a "-" sign)
- the decision to *not* make a call was appropriate (commonly noted with an "o")
- a call was made when no fault occurred (commonly noted with and "x")

On the example sheet enclosed, when the score was 3-2 in the first game, a foul was called on a pass. The rater felt the call was appropriate (noted with a "+").

Figure 5: Practical Rating Sheet Instructions

Appendix F: PAVO/USAV Rating Criteria

PAVO/USAV FIRST REFEREE CRITERIA

Here are the expanded descriptions of the elements on the rating sheet:

A. Judgment

1. Exhibits consistent judgment as follows:
   a) Maintains standards from set to set within a match.
   b) Applies the same standards for both teams.
   c) When teams have different degrees of skill, maintains consistency appropriate to both sides.
   d) Judges all players the same, regardless of whether they are setters, hitters, defensive specialists, etc.
   e) Allows appropriate length of contact for various actions (1- and 2-handed sets, passes, blocks, tips, attacks, etc.).
2. Allows appropriate latitude on sets & judges them correctly as throws or double hits.
3. Allows appropriate latitude on passes & judges them correctly as throws or double hits.
4. Allows appropriate latitude on first team contacts & judges them consistently.
5. Allows appropriate latitude on tips and attacks, & judges them consistently.
6. Allows appropriate latitude on net and block recoveries.
7. Allows appropriate latitude on blocks that are given impetus & direction; judges other illegal block actions appropriately.
8. Makes judgment based on actual contact rather than body position or technique. Avoids "automatic" calls based on a particular situation. Avoids preconceived judgments.
9. Recognizes & calls back row player faults & illegal attacks from libero sets. Shows awareness when these plays occur.
10. Calls net faults appropriately, but gives second referee time to call most nets. Allows appropriate & consistent latitude on net contacts that do not interfere with play.
11. Calls a fault if a ball touches or illegally crosses over/outside an antenna.
12. Detects & calls touches off the block.
13. Is observant for screens or potential screens & appropriately prevents. Calls screening when needed.
14. Recognizes & calls position/rotation faults on serving team.
15. Adjusts judgment to the level of competition.
16. Follows fast play at the net & adjusts to tempos within a rally.
17. Makes quick, accurate judgment calls; not influenced by ball spin, crowd or participant reactions, etc.
18. Anticipates unusual situations – antenna plays, pancakes, etc.

B. Mechanics

1. Scans benches, court, & officiating team prior to each serve.
2. Uses the correct hand signals. Signals are executed clearly & held long enough to be seen by everyone.
3. Exhibits solid knowledge of current officiating techniques.
4. Uses correct signaling sequence in appropriate tempo.
5. Whistle is clear, sharp, authoritative & audible. Varies whistle tones to differentiate match situations (for example, timeouts versus infractions versus an object on the court).
6. Has quick reaction time with the whistle.
7. Mimics signals initiated by the second referee, such as timeouts or substitutions. Mimics second referee fault signals (NCAA); repeats as a clarification when needed (USAV).

C. Positioning

1. Adjusts view appropriately:
   - Narrows focus to the plane of the net as the play moves into that area.
   - Expands view as the play broadens.
2. Sees each contact of the ball, particularly third team hits near the first referee's stand
3. Does not follow ball (too high into the air), but utilizes a lateral eye movement to adjust focus to each point of contact.
4. Quickly adjusts position between plays to obtain the best view. Maintains a stationary position when ball is being contacted.
5. Maintains proper positioning & posture.

D. Match Control

1. Inspects playing site & ensures that equipment & facilities meet accepted standards for play & for safety.
2. Oversees warm-ups or delegates warm-up supervision to second referee.
3. Follows specified match protocol; efficiently manages time.
4. Gives verbal warnings at appropriate times with an appropriate demeanor.
5. Assesses sanctions when warranted or when requested by second referee, using appropriate procedures.
6. Protects other officials from abuse by participants/spectators.
7. Demonstrates awareness of bench activity, including sub or timeout requests, second referee communications with bench personnel, substitute locations, libero exchanges, etc. Is aware of bench conduct & assists second referee with bench control if necessary.
8. Maintains awareness of on-court happenings – interference, injuries, spectators and/or media.
9. Stays alert to safety issues from arrival to departure.
10. Establishes & maintains a brisk, but unhurried match tempo.

E. Communication with Match Participants

1. Conducts pre-match meeting and coin toss in a clear, concise manner.
2. Addresses team members respectfully & maintains a dignified manner.
3. Maintains an approachable demeanor & a positive attitude.
4. Initiates conversation with captains or other team members when appropriate; interacts professionally & efficiently without causing unnecessary delay. Displays & executes superior verbal & non-verbal communication skills; avoids inappropriate body language.
5. Acknowledges concerns of coaches & captains when they disagree, but minimizes delay of the match.

F. Communication with Officiating Team

1. Fully briefs second referee on techniques & responsibilities. Instructs line judges & scorekeepers clearly & completely.
2. Displays superior verbal & non-verbal communication skills, avoids inappropriate body language.
3. Frequent eye contact with second referee & line judges.
4. Interacts well with co-officials, & accepts their assistance.
   - Encourages line judges to make their calls.
   - Allows time for scorekeepers to record events.
   - Allows the second referee to deal with his/her primary responsibilities (net, center line, antenna, etc.), & duties (requests for timeouts & substitutions or addressing concerns from the bench) when possible.

G. Professionalism

1. Wears the correct uniform that is clean & pressed. Appearance is well-groomed and professional.
2. Treats the match & participants with respect & attention. Displays dignified conduct.
3. Recovers after controversial or adverse situations & maintains appropriate tempo for the established level of play.
4. Displays a calm, decisive manner & a low-key, cordial attitude.
5. Projects confidence in oneself & the entire officiating team.
6. Displays a spirit of cooperation toward officiating team, coaches, players, event management, & the media.
7. Maintains appropriate attitude throughout match.

Revised July 2011

Figure 6: PAVO/USAV First Referee Rating Criteria

PAVO/USAV SECOND REFEREE CRITERIA

Here are the expanded descriptions of the elements on the rating sheet:

A. Judgment

1. Calls net faults; allows appropriate latitude on net contacts that do not interfere with play.
2. Recognizes & calls center line faults.
3. Recognizes & calls position faults on the receiving team.
4. Recognizes & calls back row player faults & illegal attacks from libero sets. Shows awareness when these situations occur, legally & illegally.
5. Detects & immediately indicates touches off the block after a rally.
6. Whistles antenna faults appropriately.
7. Anticipates unusual plays, including antennas, pancakes, views back row attack hits near 3-meter line, etc.
8. Adapts to level of competition.

B. Mechanics

1. Uses the correct hand signals. Signals are executed clearly & held long enough for communication.
2. Exhibits solid knowledge of current officiating techniques.
3. Uses correct signaling sequence in appropriate tempo.
4. Mimics all appropriate signals given by first referee, & is visible to first referee while signaling.
5. Whistle is clear, sharp, authoritative, & audible. Varies whistle tones to differentiate match situations (for example, timeouts versus infractions versus an object on the court).
6. Has quick reaction time with the whistle.
7. Scans benches, court, & first referee prior to each serve & at the end of each play. Shows awareness of libero replacements (and libero serving for NCAA).

C. Positioning/Focus

1. Obtains a clear view of the receiving team prior to each serve, but is positioned so that peripheral vision can detect activity from the serving team's bench.
2. Focuses attention on the blocking side of the net during play.
3. Position is 3-6 feet from the net post as facilities allow.
4. Transitions immediately on service contact. Transitions quickly to the blocking side during rallies, but is stationary at ball contact.
5. After moving to blocking side, is able to further adjust position to see the play develop or assist first referee.
6. Establishes a position that enables view of both blockers & attackers through the net, shoulders/feet slightly angled to open body to attacker's side.
7. Is able to follow fast play at the net & adjust to varying tempos during a rally.
8. To give assistance or at the end of each rally, steps to one side of the net support & is clearly visible to first referee.
9. Broadens & narrows focus appropriately during play.

D. Match Control

1. Obtains line-ups at specified times, & uses preventive measures to correct potential problems. Affirms uniform & player equipment legality. Checks rosters (if applicable). Controls intervals between sets.
2. Efficiently controls, oversees & times warm-up periods.
3. Checks lineups on the court prior to each set & identifies captains for first referee. Authorizes libero entry before set.
4. Makes decisions quickly & accurately.
5. Facilitates communication among the benches, line judges, scorekeepers, first referee, game management, & media.
6. Observes & reacts appropriately to occurrences during play & between rallies.
7. Responds quickly to timeout & substitution requests from the bench while minimizing delays to the match.
8. Facilitates substitution from position near the score table & authorizes entry. Communicates status of substitution process to coach, scorekeepers, & first referee.
9. Allows appropriate amount of time & ends timeout period with a warning whistle (if appropriate) & a final whistle. Informs first referee and coach/captain of the number of timeouts taken by each team.
10. Ensures that team members remain in the bench or warm-up areas as appropriate. Is aware of bench conduct. Appropriately deals with questions & concerns from the bench. Addresses coaches' concerns out of the front zone, and does not delay the match.
11. Gives verbal warnings at appropriate times with an appropriate demeanor.
12. Requests sanctions from first referee when warranted.
13. Stays alert to safety issues from arrival to departure.
14. Protects officiating team from abuse by players, coaches, & spectators.
15. Helps first referee establish & maintain appropriate match tempo. Supervises floor wiping by players, minimizing delay.

E. Communication with Match Participants

1. Addresses players/coaches respectfully, in a dignified way.
2. Maintains an approachable demeanor & positive attitude.
3. Initiates conversation with captains or other team members when appropriate; interacts professionally & efficiently without allowing unnecessary delay; displays & executes superior verbal & non-verbal communication skills.
4. Acknowledges concerns of coaches & captains when they disagree, but minimizes delay of the match.

F. Communication with Officiating Team

1. Reviews responsibilities & techniques with first referee.
2. Ensures lineup is correctly entered on scoresheet. Reviews duties with other score table staff. Communicates clearly with scorekeepers during the match.
3. Displays superior verbal & non-verbal communication skills with other officials; avoids inappropriate body language.
4. Makes eye contact with first referee at the end of rallies, & makes eye contact with other officials as needed.
5. Offers discreet assistance to the first referee on calls where the first referee may not have a clear view of the play.
6. Encourages co-officials after difficult decisions.
7. Is alert to addressing questions or concerns from line judges. Clarifies calls made by line judges when needed.
8. Coordinates the work of speed wipers & ball crew.

G. Professionalism

1. Wears a correct uniform that is clean & pressed; is well-groomed & professional in appearance.
2. Treats the match & participants with respect & attention. Dignified conduct.
3. Recovers after controversial or adverse situations & maintains appropriate tempo for the established level of play; handles unusual situations well.
4. Displays calm, but decisive manner & a low-key but cordial attitude.
5. Projects confidence in oneself & in the performance of the entire officiating team.
6. Displays a spirit of cooperation toward officiating team, coaches, players, event management, & the media.
7. Maintains appropriate attitude throughout the match.

Revised July 2011

Figure 7: PAVO/USAV Second Referee Rating Criteria

REFERENCES

REFERENCES

Anderson, K. J., & Pierce, D. A. (2009). Officiating bias: The effect of foul differential on foul calls in NCAA basketball. *Journal of Sports Sciences, 27*(7), 687–694.

Anshel, M. H. (1995). Development of a rating scale for determining competence in basketball referees: Implications for sport psychology. *The Sport Psychologist, 9*, 4–28.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191.

Bandura, A. (1982). Self-efficacy in human agency. *American Psychology, 37,* 122-147.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. (2001). Social-cognitive theory: An agentic perspective. *Annual Review of Psychology, 52,* 1-26.

Boyko, R. H., Boyko, A. R., & Boyko, M. G. (2007). Referee bias contributes to home advantage in English Premiership football. *Journal of Sports Sciences, 25*(11), 1185–1194.

Catteeuw, P., Helsen, W., Gilis, B., & Wagemans, J. (2009). Decision-making skills, role specificity, and deliberate practice in association football refereeing. *Journal of Sports Sciences, 27*(11), 1125–1136.

Dosseville, F., Laborde, S., Raab, M., & others. (2011). Contextual and personal motor experience effects in judo referees' decisions. *Sport Psychologist, 25*(1), 67.

Dowrick, P.W., & Dove, C. (1980). The use of modeling to improve the swimming performance of spina bifida children. *Journal of Applied Behavior Analysis, 13,* 51-56.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41,* 1149-1160.

Feltz, D.L., Short, S.E., & Sullivan, P.J. (2008). *Self-efficacy in sport*. Champaign, IL: Human Kinetics.

Guillén, F., & Feltz, D. L. (2011). A conceptual model of referee efficacy. *Frontiers in Psychology, 2,* 25.

Huck, S. (2004). *Reading statistics and research (4th ed.)*. Upper Saddle River, NJ: Ally & Bacon.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.

Lehman, D. R., & Reifman, A. (2001). Spectator influence on basketball officiating. *The Journal of Social Psychology*, *127*(6), 673–675.

Lopez, M. J., & Snyder, K. (2013). Biased impartiality among National Hockey League referees. *International Journal of Sport Finance*, *8*, 208–223.

MacMahon, C., Helsen, W. F., Starkes, J. L., & Weston, M. (2007). Decision-making skills and deliberate practice in elite association football referees. *Journal of Sports Sciences*, *25*(1), 65–78.

Maddux, J.E. (1995). Self-efficacy theory: An introduction. In J.E. Maddux (Ed.), *Self-efficacy, adaptation, and adjustment: Theory, research, and application*. New York: Plenum Press.

McInman, A. D. (1997). *Where are all the sport psychology umpire studies?* Presented at the 32nd Annual Conference of the Australian Psychological Society, Cairns, Australia.

Moritz, S.E., Feltz, D.L., Fahrbach, K.R., & Mack, D.E. (2000). The relation of self-efficacy measures to sport performance: A meta-analytical review. *Research Quarterly for Exercise and Sport, 71*, 280-294.

Myers, N. D., Feltz, D. L., Guillén, F., Dithurbide, L., & others. (2012). Development of, and initial validity evidence for, the referee self-efficacy scale: A multistudy report. *Journal of Sport and Exercise Psychology*, *34*(6), 737.

Nunnaly, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

Pizzera, A., & Raab, M. (2012a). Does motor or visual experience enhance the detection of deceptive movements in football? *International Journal of Sports Science and Coaching*, *7*(2), 269–284.

Pizzera, A., & Raab, M. (2012b). Perceptual judgments of sports officials are influenced by their motor and visual experience. *Journal of Applied Sport Psychology*, *24*(1), 59–72.

Plessner, H., & Betsch, T. (2001). Sequential effects in important referee decisions: The case of penalties in soccer. *Journal of Sport and Exercise Psychology*, *23*, 254–259.

Souchon, N., Cabagno, G., Rascle, O., Traclet, A., Dosseville, F., & Maio, G. R. (2009). Referees' decision making about transgressions: the influence of player gender at the highest national level. *Psychology of Women Quarterly*, *33*(4), 445–452.

Souchon, N., Cabagno, G., Traclet, A., Dosseville, F., Livingstone, A., Jones, M., & Maio, G. R. (2010). Referees' decision-making and player gender: the moderating role of the type of situation. *Journal of Applied Sport Psychology*, *22*(1), 1–16.

Souchon, N., Cabagno, G., Traclet, A., Trouilloud, D., & Maio, G. (2009). Referees' use of heuristics: The moderating impact of standard of competition. *Journal of Sports Sciences*, *27*(7), 695–700.

Souchon, N., Coulomb-Cabagno, G., Traclet, A., & Rascle, O. (2004). Referees' decision making in handball and transgressive behaviors: Influence of stereotypes about gender of players? *Sex Roles, 51*(7-8), 445–453.

Trudel, P., Cote, J., & Sylvestre, F. (1996). Systematic Observation of Ice Hockey Referees During Games. *Journal of Sport Behavior*, *19*(1), 66–81.

Unkelbach, C., Memmert, D., & others. (2008). Game management, context effects, and calibration: The case of yellow cards in soccer. *Journal of Sport and Exercise Psychology*, *30*(1), 95.

van Quaquebeke, N., & Giessner, S. R. (2010). How embodied cognitions affect judgments: Height-related attribution bias in football foul calls. *Journal of Sport & Exercise Psychology*, *32*, 3–22.

Vealey, R. S., Hayashi, S. W., Garner-Holman, M., & Giacobbi, P. (1998). Sources of sport-confidence: conceptualization and instrument development. *Journal of Sport & Exercise Psychology*, *20*, 54–80.

Wagner-Egger, P., Gygax, P., & Ribordy, F. (2012). Racism in soccer? Perception of challenges of black and white players by white referees, soccer players, and fans. *Perceptual and Motor Skills*, *114*(1), 275–289.

Weinberg, R., Gould, D., & Jackson, A. (1979). Expectations and performance: An empirical test of Bandura's self-efficacy theory. *Journal of Sport Psychology*, *3*, 345-354.