

CALIBRATION AND ACHIEVEMENT GOALS IN COLLEGE VOLLEYBALL

by

Fred Chao  
A Thesis  
Submitted to the  
Graduate Faculty  
of  
George Mason University  
in Partial Fulfillment of  
The Requirements for the Degree  
of  
Master of Science  
Educational Psychology

Committee:

*Angela D Miller*

Chair

*Michelle M. Full*

*Greg E. Kelly*

*Anastasia Kitzon*

Program Coordinator

*Martin Ford*

Dean, College of Education  
and Human Development

Date:

*April 22, 2014*

Spring Semester 2014  
George Mason University  
Fairfax, VA

Calibration And Achievement Goals In College Volleyball

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at George Mason University

By

Fred Chao  
Bachelors of Science  
George Mason University, 1994

Chair: Angela Miller, Assistant Professor  
Department of Educational Psychology

Spring Semester 2014  
George Mason University  
Fairfax, VA

## Dedication

This is dedicated to my wife, Dana, whose love and support inspires me to be the best husband and father I can be; to our new daughter, Keira Makenzi, who has given me purpose and perspective; and to our dog, Mason, who has kept me company through the late nights of writing.

## Acknowledgements

I am truly grateful to all of the wonderful people who are part of the Educational Psychology department at George Mason. My thanks to Dr. Gorrell, for providing my first course and getting me hooked; to Dr. Kalbfleisch, who taught me to read research; to Dr. Sheridan, who reminded me that people complicated and interesting; to Dr. Kitsantas, who provided me with the encouragement to pursue and complete my degree; and to Dr. Bland, whose unwavering belief in me is humbling. To the members of my committee, I am forever in debt for your willingness to take a special interest in my pursuit of learning. To Dr. Buehl, who motivated me to be a better coach; to Dr. Kelly, who constantly reminded me that perspective is everything. Finally, I would like to thank Dr. Miller, whose skill as an instructor and a guide has impressed upon me a standard to which I will aspire, and for her ability to connect my world of collegiate athletics with scholarship.

## Table Of Contents

|  |     |
|--|-----|
| List of Tables.....                      | v   |
| List of Figures.....                     | vi  |
| Abstract.....                            | vii |
| 1. Introduction.....                     | 1   |
| 2. Review of Literature .....            | 4   |
| 3. Purpose.....                          | 29  |
| 4. Methods.....                          | 31  |
| 5. Results.....                          | 39  |
| 6. Discussion .....                      | 59  |
| 7. Implications.....                     | 68  |
| 8. Limitations and Future Research ..... | 72  |
| 9. Conclusion .....                      | 76  |
| 10. Appendices.....                      | 77  |
| 11. References.....                      | 89  |
| 12. Biography .....                      | 97  |

## List Of Tables

| Table  | Page |
|--|------|
| 1. Demographic Information.....                | 40   |
| 2. Summary of Measures Descriptives .....      | 42   |
| 3. Summary of Correlations.....                | 45   |
| 4. Bases of Perceived Opponent Difficulty..... | 56   |

## List Of Figures

| Figure   | Page |
|--|------|
| 1. Scatterplot of confidence versus accuracy (as indexed by discrepancy) .....               | 49   |
| 2. Scatterplot of confidence intervals and discrepancy with frequencies .....                | 50   |
| 3. Scatterplot of confidence intervals versus discrepancy with respect to performance.       | 52   |
| 4. Scatterplot of mean confidence versus mean discrepancy for three performance groups ..... | 54   |
| 5. Average bias for each player ranked by performance .....                                  | 55   |

## Abstract

### CALIBRATION AND ACHIEVEMENT GOALS IN COLLEGE VOLLEYBALL

Fred Chao, MS

George Mason University, 2014

Chair: Dr. Angela Miller

The purpose of this study was to quantitatively explore the relationship between calibration and achievement goals to attack performance in male, collegiate volleyball players. Three primary exploratory research questions were examined: 1) what patterns of goal orientation exist, 2) what are the relationships among calibration accuracy, calibration confidence, experience, and performance, and 3) what patterns of calibration exist? A study-specific questionnaire was used to measure achievement goal orientation, experience, and expected performance. Performance was measured from match data collected on three separate occasions throughout the course of a single-day volleyball tournament. Preliminary data analysis supported a 2 x 2 achievement goal framework and multiple goals theory. Cluster analysis yielded two distinct goal profiles: *high-all*, and *high-approach*. Independent t test revealed no relationship between the profiles and either experience or performance. Correlational analyses showed that high-mastery oriented players were more accurate in their estimates of performance (better calibrated), while



high-performance oriented players were more confident in their estimates. Finally, graphical analysis showed a slight overall tendency for overconfidence in estimates for performance, with higher-performing players showing more underconfidence and lower-performing players showing more overconfidence. Implications for practice and future research were discussed.

## CALIBRATION AND ACHIEVEMENT GOALS IN COLLEGE VOLLEYBALL

Successful athletic teams are externally judged as those that win more than lose. As a collegiate men's volleyball coach, I am keenly interested in seeking methods to increase current performance levels in order to win more matches. However, underpinning my declared interest are two assumptions that 1) current performance levels are known, and 2) motivation for improvement exist. In educational research terms, I am interested in performance calibration and achievement goal theory.

The degree to which athletes can accurately calibrate their ability along with their achievement goal orientation can effect development and, ultimately, competitive performance. For an athlete to determine whether improvement has been made, it would seem logical to accurately estimate current ability. Based on personal experience, however, athletes vary in their ability to make accurate estimations. Consider, for example, the freshmen athlete who is unaware of his current ability, having just transitioned from a lower level of high school age competition to National Collegiate Athletic Association (NCAA) Division I competition, which is the highest level of indoor volleyball league in the United States. Thus, part of the adjustment for the freshmen is a calibration of perceived ability.

Historically, the construct of calibration is rooted in research of human decision-making under conditions of uncertainty, which looked to study people's confidence in

their judgments and predictions (Fischhoff, Slovic, & Lichtenstein, 1977). The construct has since been sparingly studied in the sport domain, where calibration has been defined as “the accuracy with which one can rate or predict one’s own performance” (Fogarty & Else, 2005, p. 42). If a player and coach disagree on current ability level, it will be difficult for the player to accept his role as defined by the coach. Player development may also be affected by calibration accuracy. A player who is highly overconfident might not respond to corrective feedback or training accuracy. Unfortunately, research literature regarding calibration in sport is limited. A goal of this study is to examine the calibration accuracy of competitive volleyball players.

Competitive athletes are also concerned with developing *and* demonstrating ability. Achievement goal theory (AGT) divides these pursuits as mastery and performance goals, respectively. AGT research in classroom settings suggests that pursuing mastery goals, where achievement is self-referenced and based on improvement, is more efficacious for learning and achievement than performance goals (Ames, 1992; Elliott & Dweck, 1988; Nicholls, 1984). Athletes, on the other hand, compete against opponents on a regular basis, and may be more inclined to judge achievement relative to others, thus promoting a performance goal orientation. However, literature specifically addressing the achievement goal profile of competitive athletes is also limited. Another goal of this study is to examine the achievement goal orientations of competitive volleyball players.

Without accurate performance self-calibration, players’ rely solely on the ability of a coach to gauge improvement, rather than take responsibility for their development.

However, even if a player is well-calibrated for performance, improvement does not necessarily follow. Performance and improvement are influenced by the motivation to pursue mastery or learning goals. Therefore, given the implications for performance and learning in competitive sport, this study aims to extend the research on calibration and achievement goals in collegiate athletes.

By using quantitative approaches to examine calibration and achievement goals in collegiate athletes, we may identify calibration and achievement goal characteristics of this unique population that predict players' ability to develop and perform. With this information, players, coaches, and sports psychologists may be better equipped to structure and implement player development plans, manage expectations, and ultimately improve performance. Additionally, researchers can create or modify current theories of calibration and achievement goals.

## **Review of Literature**

Existing literature on calibration and AGT is primarily focused on students and classroom learning. Initial research reveals a large number of studies relating calibration or AGT as constituent components of various other constructs and theories, including, but not limited to, metacognition, self-concept, self-regulation, and self-efficacy. However, even a brief examination of the multiple applications of calibration and AGT is beyond the scope and intent of the present study. Therefore, only those studies that inform a general understanding of the construct or application to the sport specific domain will be reviewed. Calibration will be addressed first, followed by AGT.

### **Calibration**

In an introduction to a special publication on calibration, Alexander (2013) offers calibration as “the distance between perceived and demonstrated levels of understanding, capability, competence, or preparedness” (p. 1). Simply, how close do people’s judgments actually match reality? The construct of calibration is rooted in research of human decision-making under conditions of uncertainty, which looked to study people’s confidence in their judgments and predictions (Lichtenstein, Fischhoff, & Phillips, 1981). Calibration research has followed a general paradigm whereby participants are asked to answer a question of general or specific knowledge, and then asked to indicate a confidence rating for the answer (Lichtenstein et al., 1981). After a series of questions,

the data is analyzed for performance accuracy and confidence. Performance accuracy is calculated as the percentage of correct answers. One common method for determining calibration is by calculating the difference between mean confidence rating and percent correct answers. For example, if over a series of questions, the mean confidence rating is 70% and the percentage of correct answers is 70%, the participant would be well-calibrated, whereas 50% correct answers would indicate overconfidence, and 90% correct answers would indicate under-confidence. A second commonly used method for determining more precise variations in calibration levels is the comparison between the percent of answers correct for all questions assigned a given confidence rating. As an example, for all questions assigned a confidence rating of 70%, the percent of correct responses should be 70% for perfect calibration. Correct response percentage of 50% would indicate overconfidence, and 90% would indicate underconfidence. The comparison can be repeated over all confidence ratings, and patterns of accuracy can be examined with respect to the various confidence levels. While overall percent correct responses compared to mean confidence rating can offer a global indication of calibration, the ability to look at calibration for a given confidence rating is useful to explore variations in confidence levels.

### **Patterns of Calibration**

Throughout much of the calibration literature, patterns have emerged that have become the focus of further research. One such pattern is the finding that people tend to overestimate their ability. That is, people are generally overconfident. Another pattern is

the hard-easy effect, whereby people generally show overconfidence for difficult tasks and underconfidence for easy tasks. Both of these patterns will be expounded upon.

**Overconfidence.** One of the early patterns to emerge in calibration research is the tendency for people to be overconfident in their judgments (Fischhoff et al., 1977; Koriat, Lichtenstein, & Fischhoff, 1980; Lichtenstein et al., 1981; Lichtenstein & Fischhoff, 1977). In a series of five studies using paid college students participating in answering questions of general knowledge, Fischhoff and Lichtenstein (1977) examined the pattern of overconfidence. In the first experiment, the format for the questions varied and included open-ended (43 questions), one-alternative (75 questions), and two variations of two-alternative responses (75 and 50 questions). After each question, participants were asked to write their probability of answering the question correctly using a typical probability scale of 0.00 – 1.00. It was found that at the extremes of the scale (absolute uncertainty and absolute certainty), subjects were correct between 20-30%. The remainder of the experiments confined the question format to two alternatives. Subjects were instructed to select the correct answer, and then indicate a confidence level. However, unlike experiment 1, subjects were asked to indicate confidence in terms of odds, rather than probability. The researchers rationale for making the change is that the probabilities, particularly at the high end of the scale, lacked sensitivity. The authors explain that typical probability responses of .90, .95, and 1.00 correspond to odds of 9:1, 19:1, and  $\infty$ :1, respectively (Fischhoff et al., 1977), which lack the discriminative sensitivity that may be required for more accurate confidence ratings. In experiment 2, 51% of the total number of judgments was at 50:1 or higher, correlating to probabilities

of 98% or higher. The % correct answers for the range of odds between 50:1 to 1,000,000:1 was 68-90%, indicating overconfidence, particularly in high certainty judgments. The remaining experiments showed similar patterns, although a more careful explanation in experiment 3 resulted in fewer judgments at extreme odds (50:1 or greater). In the final two experiments, the researchers challenged the subjects to a hypothetical gambling game based on odds at or greater than 50:1. For every wrong answer given by a subject at those odds, the subject was to pay the researcher \$1. For each wrong answer, the researcher would select from a bag of 100 white and 2 red poker chips, which is equivalent to 50:1 odds. For every red chip that is selected, the researcher would pay the subject \$1. Because the actual ratio of incorrect to correct answers at these odds were found to average 1:8, rather than 1:50, the subjects would have lost money to the researchers. The gambling game highlighted the findings from all five experiments that people are overconfident in their probability judgments across all types of question formats, or they do not understand odds. This pattern of overconfidence has become a target of study throughout calibration research, which will be discussed later in this review.

**Hard-easy effect.** The hard-easy effect is, simply, the trend in calibration studies to show overconfidence for hard questions and underconfidence for easy questions (Koriat, Sheffer, & Ma'ayan, 2002; Lichtenstein et al., 1981). Hard items are classified as such because fewer participants answer them correctly, while most participants answer easy items correctly. In comparison to the line of perfect calibration whereby confidence ratings are matched with proportion correct, the relative proportion of correct answers



falls below this line for hard items indicating overconfidence, and above this line for easy items indicating underconfidence.

While both the overconfidence and the hard-easy effect has been found throughout calibration research (see review by Keren, 1991), the actuality of the effects has become a topic of debate (Juslin, Winman, & Olsson, 2000). Specific details of the argument against these common patterns are beyond the scope of the present review. However, in a quantitative review of the data from available studies at the time, Juslin et al. (2000) argues that the hard-easy effect and overconfidence tendency is an artifact of experimental method, including scale-end effects, linear dependency, and regression effects, rather than informational processing biases. As such, attributions of calibration patterns to characteristics of the assessor (Keren, 1991) cannot be determined. Therefore, although general patterns of calibration accuracy have been demonstrated, the causes for the patterns are not clearly identified. Perhaps the patterns appear because calibration is related to the different cognitive processes between experts and novices, or perhaps because they are merely artifacts of method or task difficulty.

Determining if patterns exist in competitive athletes could inform more effective coaching practices. For example, players who demonstrate the overconfidence pattern would be less receptive to change, or less “coachable”. On the other hand, even the armchair coach knows that a high degree of confidence is needed to withstand the stresses of competition. The hard-easy pattern may indicate that calibration accuracy and confidence are not related, and one aspect may be important for learning and development, while the other for performance.

## **Factors That Influence Calibration**

In order to identify potential causes of the general patterns, the components of calibration must first be identified. As such, numerous studies have attempted to influence calibration accuracy by looking factors associated with task difficulty, expertise, reasoning, feedback, and practice.

**Task difficulty.** Methodological arguments against the hard-easy effect are echoed in conclusions drawn by Burson, Larrick, and Klayman (2006) that both skilled and unskilled performers show equivalent levels of miscalibration. The researchers looked at calibration and its relationship with task difficulty. Additionally, they looked at people's judgments of performance relative to their peers, rather than just their own performance. Subjects were students recruited through ads at the University of Chicago and paid for their participation. Three studies were performed, one was a trivia quiz with a range of difficulty, the second was a trivia quiz with questions selected to be more moderate to difficult, and the third was a word game that was also differentiated by difficulty. After taking a quiz, participants indicated their estimates of performance, their percentile rank relative to peers, their perceived difficulty of the task for themselves, and for their peers. The authors found both skilled and unskilled performers (defined by performance on the tasks) to be inaccurate in their estimates. The researchers discussed the phenomenon of most people to judge themselves as "better than average" for easy tasks and "worse than average" for difficult tasks. That is, the skilled do not possess greater judgment abilities; rather, because they performed better than average in easy tasks, they are accurate in their estimates. Likewise, because the unskilled performed

worse than average in difficult tasks, they were also accurate in their estimates. For clarity, the authors explain that in the instance of difficult tasks:

We do not suggest that poor performers are actually more perceptive than high performers in these tasks. Rather, in a task in which everyone is biased toward believing their performance is poor, those whose performance truly is poor will appear to be right. (Burson et al., 2006, p. 65).

The reciprocal statement can be made regarding easy tasks; judgments biased toward high performance are only accurate for those whose are actually correct. Thus, the authors propose a model of “noise plus bias” to account for people’s alleged miscalibration levels. Poor judgments of ability on a task are function of random task variation (noise) and poor feedback about performance plus the phenomena of judgments being correct only as a consequence of performance (bias). Further, the appearance of these calibration patterns result as “an accident of task difficulty, not an indication of greater awareness” (Burson et al., 2006, p. 67).

Despite the methodological issues that plague calibration research, task difficulty is of clear concern. It is a variable that has been manipulated in the experimental research considered thus far, and is the source of the controversial patterns of overconfidence and hard-easy effects previously discussed.

**Expertise and performance.** Another factor that may moderate calibration related to task difficulty is expertise. A task that is relatively difficult for a novice may be relatively easy for an expert. Findings from studies looking at expertise as a factor have been mixed (Lichtenstein et al., 1981). Among a sample of nine doctors (expertise

implied), Christensen-Szalanski and Bushyhead (1981) found that, as a group, the physicians were significantly overconfident for correctly assigning a diagnosis of pneumonia. On the other hand, expert card game players showed more accurate calibration than novices (Keren, 1987). This study directly attempted to answer the question as to whether improvement in skill implies improved calibration for a given task difficulty level. Comparing 8-pairs of expert and 14-pairs of amateur bridge players (based on membership to a bridge-specific club as well as participation in national and international tournaments), the researchers found that amateurs were more poorly calibrated, and also demonstrated overconfidence. Additionally, expert players actually demonstrated a general level of underconfidence. In the present study, because the subjects all participate at a similar level (collegiate), and because there is often overlap of abilities among NCAA Division 1, 2, and 3 players, the same distinction of expertise by level of club membership or level of competitive tournament play cannot be made. The distinction could be made if the sample included professional players in contrast to collegiate players. It can be argued, however, that better actual performance is indicative of higher expertise than poorer performance. As such, performance will be used in place of expertise for the present study.

**Practice.** Practice has been shown to reduce the overconfidence bias. Koriath et al. (2002) used multiple study-test cycles on Hebrew-speaking undergraduates from the University of Haifa to look at practice effects on confidence judgments. In the first experiment, 20 subjects were displayed words from a 40-word list of common Hebrew words using a computer. After presentation of a word, they were asked for judgment of

learning rating on a scale of 0%-100%, homologous to confidence, that subjects would be able to free recall the word on the following test. The cycle of study-test was repeated four times, and results confirmed an underconfidence-with-practice effect. That is, with each successive study-test cycle, overconfidence levels reduced and subjects indicated more underconfidence in their judgments. Similar results were reported for the second experiment, using a similar group of subjects but with a different task. This time, subjects were shown a fictitious word that was associated with a physical action. Probability estimates to recall the associated action for each word were then inputted. Again, the researchers were able to decrease overconfidence.

**Feedback.** The pattern of general overconfidence was the target for studies using feedback about performance accuracy in an attempt to increase calibration (Arkes, Christensen, Lai, & Blumer, 1987; Koriatic et al., 1980; Lichtenstein & Fischhoff, 1980). In a study using 58 undergraduate students, researchers looked at feedback as an attempt to reduce overconfidence (Arkes et al., 1987). Questions of knowledge of history, geography, and sports were used. Though all of the subjects received five practice questions of similar difficulty, half were given questions that appeared difficult, while the other half were given questions that appeared easy, but were just as difficult. The difference between the types of questions was based on more recognizable topics, though the percent of correct answers for both types of questions were similar. After both groups answered the initial five questions, half of each group was given feedback on the accuracy of their answers, while the other half received no feedback. All subjects answered an additional 30 questions, with confidence ratings for each. Results showed

that, as expected, more overconfidence was found in the group who answered the easy-appearing questions compared to the difficult-appearing questions. Out of the groups to receive feedback, both the easy and difficult groups reduced their confidence, with a more pronounced underconfidence pattern emerging in the easy/feedback group.

**Making reasons salient.** Overconfidence was also reduced in a study in which subjects were asked to give reasons for selecting answers (Koriat et al., 1980). In the first experiment, 73 paid volunteers who answered an ad at the University of Oregon answered questions of general knowledge under two separate conditions. The control condition was the typical method of choosing the correct alternative and indicating a confidence rating. In the reasons condition, subjects were instructed to fill out a 2 (alternative a or b) x 2 (reasons for and reasons against) matrix, then rate each reason on a scale according to strength of reason. Finally, an answer was chosen and a confidence rating was indicated. For each condition, subjects answered 10 questions which were sampled from a pool of 150 questions of general knowledge used by Lichtenstein & Fischhoff, (1977). Overconfidence was reduced in the reasons condition while proportion of correct answers increased, thus improving accuracy.

The second experiment in the same study hypothesized that contradictory reasons would improve calibration. Researchers simplified the experimental conditions into three categories: reasons to support, reasons to contradict, and both. Subjects were also paid, volunteer respondents of an ad at the university. Of the 200 subjects, 66 were in the support condition, 66 in the contradict condition, and 68 in both conditions, though the manner of placement was not specified. The procedure for each condition was similar

Experiment 1 except that, instead of the 2 x 2 matrix, only reasons corresponding to the assigned group was to be given and no strengths of reason were required. In particular, for the contradictory reasons condition, subjects were asked to give the best contradicting reason for selecting an answer (or the best contradicting reason for not selecting an alternative). Researchers found support to their hypothesis and results showed that contradictory reasons for selecting to be more effective for reducing overconfidence than the support condition or the both condition. The researchers discussed the improvement in terms of making contradictory evidence more salient as being the reason for a reduction in overconfidence.

### **Calibration in Sport**

To this point, the present review of literature has been concerned with studies that inform the basic tenets of calibration research, the existing patterns of calibration, and the factors that have been examined to influence calibration. Attention now turns to calibration research in the domain of sports. While physical education in research literature is often studied and involves sport activities; this is not the pertinent domain. The age, goals, motivations, intensity, and performance levels of competitive, collegiate athletes are vastly different from physical education class. Given this distinction, existing research on calibration in competitive sport is limited.

**Golf.** In a study of 54 male golfers, Fogarty and Else (2005) set out to investigate overconfidence in male golfers using the calibration paradigm. Participants ranged in age from 13 to 75, were recruited through personal contact, and included a range of expertise as determined by handicap level. The tasks used in this study consisted of two

questionnaires assessing self-confidence in putting and chipping skill, two performance tasks of putting and chipping, and a golf-knowledge test. The calibration paradigm was applied to the paper-and-pencil test of golf-knowledge. After each question, a confidence rating was given. Scores were converted to percentage correct, and mean confidence rating was calculated. The researchers found a 14% difference between mean accuracy scores and mean confidence rating, and their hypothesis of the tendency to be overconfident in the golf-knowledge test was supported.

For the skill tests, participants were asked to estimate their number of successful putts or chips out of 20 before performing the task. The percentage of actual successful shots was subtracted from the percentage of estimated successful shots, and the bias was the difference between the obtained and estimated scores. It was hypothesized that levels of calibration would be good for these sport-specific tasks. The researchers also hypothesized that low-handicap (more skilled) golfers would be better calibrated than the high-handicap (less skilled) golfers. Finally, the researchers used results from previous trials as feedback, and hypothesized that calibration would improve on the basis of this feedback (it could be argued that this is practice rather than feedback). Using multi-factorial ANOVA, none of these relationships reached significance, although the feedback condition just missed,  $F(2, 51) = 2.89, p = .06$ .

**Tennis.** In a study of tennis players, Fogarty and Ross used calibration to “describe the extent which people are accurate in [their] self-assessments” (2007, p. 148). Their research in the sport of tennis highlights the importance of being “well-calibrated” as “the ability to be realistic when rating previous performance and making future



probability judgments” (p. 148). Past performance guides beliefs about present abilities, which influences predictions about future performances. Participants in the study were 64 (41 male, 23 female) tennis players with an age range of 14-48 years, and were recruited by personal contact and selected based on variation in expertise and gender. Current and past professional players were deemed experts, while social adult and tournament standard junior players were deemed non-experts. The authors discussed the overconfident pattern that was explained earlier in this review, and hypothesized that players would be overconfident in all experimental tasks. The second hypothesis was that expertise would correlate with better calibration. The third hypothesis was that feedback would improve calibration. Five tasks were used as tests. The first three were paper-and-pencil, multiple-choice tests about tennis rules, tennis general knowledge, and tennis technique. After each question, participants were asked to indicate how confident they were about the correctness of their answer as a percentage on a 4-choice scale (25%, 50%, 75%, 100%). This gave the researcher three scores: correct answers converted to percentages, confidence rating as a percentage, and a bias score, which was the difference between the predicted and obtained scores. However, the methods section did not mention at what point the predicted scores were obtained. Reference to Fogarty’s previous research points to the bias score as the difference between the average confidence rating and the average correct answers. As in the previously mentioned study, positive bias scores were said to suggest over-confidence and negative bias scores were said to suggest under-confidence. Overconfidence was found in the bias scores for the

questions about tennis rules and general tennis knowledge, but the questions of tennis technique showed good calibration.

The two remaining tasks were performance tests on the skill of serving. The first performance test was to serve tennis balls at a specified target area. The second performance test reduced the target area by 50%. In each of these performance tasks, participants were asked to estimate how many serves they would make in the target zone out of 10 attempts. This gave the researchers three scores as well: a predictive serving estimate converted to a percentage, an actual serving score converted to percentage, and a bias – the difference between estimated and obtained scores. Repeated-measures ANOVA using calibration, task, and trial showed no significant interactions. However, an effect for calibration was shown, indicating overconfidence across the serving tasks.

An effect of expertise on calibration was found for the second, more difficult serving task, but not for the easier one. Thus, the relationship between expertise and calibration showed mixed results. The researchers offered that expert tennis players might be more familiar with aiming for the smaller target than novice players.

Finally, the authors again defined previous trials as feedback, rather than practice, and found that for the easier serving task, 48.28% of the participants improved calibration from Trial 1 to Trial 2. Of the remaining participants, 24.14% became worse, and 27.59% stayed the same. For the second, more difficult serving task, only 36.21% improved, while 34.48% became worse, and 29.31% remained the same. Thus, the authors were able to partially support a general overconfidence in all five tasks, but were not able to discern any significant findings for the expertise and feedback hypotheses.

Based on the findings of the various research studies reviewed to this point, calibration research is untidy at best. Even the most consistent patterns of overconfidence and the hard-easy effect may not actually exist. The research in competitive sport is limited and has not demonstrated any discernable patterns, except for overconfidence (Fogarty & Else, 2005; Fogarty & Ross, 2007). Though expertise and feedback failed to relate to calibration changes in these two studies, they were shown to improve calibration in the cognitive domain (Arkes et al., 1987; Keren, 1987; Lichtenstein & Fischhoff, 1980). Thus, further investigation into the influences of calibration is required. Discerning among the factors that influence calibration has not been clear, either. Calibration differences between trials could be considered practice or feedback, and task difficulty can be a function of expertise, rather than of the task.

As mentioned earlier, the difference in population between competitive collegiate student-athletes and others that have been studied is significant. By virtue of the requirement to compete, this group is regularly assessed on a more public stage. As such, they face the concurrent tasks of improving ability and demonstrating it. College volleyball teams are often comprised of the better performers of various regions. When placed together, their norm group has drastically changed. Calibration, or recalibration, may provide a foundation for establishing roles, managing expectations, and player development. All of which are critical ingredients to success in athletics. Thus, understanding the calibration patterns exhibited by this group and continuing the search for factors influencing calibration accuracy are important areas of research.

### **Achievement Goal Theory**

Distinguishing between improvement and results relates to an individual's motivation to pursue different goals. AGT posits that competence-related behavior is goal directed, and these goals are divided into those that promote adaptive motivational patterns of behavior and those that promote maladaptive patterns (Ames, 1992; Dweck, 1986). A variety of terminology has been used throughout the literature; task, mastery, and learning all refer to goals that aim to develop competence, while ego and performance refer to goals that aim to demonstrate competence (Dweck, 1986; Elliott & Dweck, 1988; Nicholls, 1984). In the review that follows, terminology that was used by the authors of the particular study will be maintained. However, mastery and performance will be adhered to for presentation of the present study, subsequent to the review. At the heart of the theory is the belief that achievement is either a function of effort or ability, and the motivations for engaging in a task are to increase competence relative to the task or to oneself, or to demonstrate competence relative to others (Ames, 1992; Dweck, 1986; Nicholls, 1984).

**Effort versus ability.** Children's ability to distinguish between effort and ability develop over time, and the interrelations between them also depend on the relative complexity of the task. In a study with eight boys and eight girls of each age from 5-13, Nicholls (1978) looked at developmental changes in achievement-related behavior. Films of pairs of children modeling two different levels of math problem-solving behaviors were shown to the subjects. One model was shown to be reading or writing for the entire 90 seconds, while the other spent only 40 seconds engaged in similar behavior and the other 50 seconds fiddling or looking around the room. After viewing the film, each

subject was told one of three different result conditions for the model pair. One condition was that both model children achieved the same high score of 10 out of 10. The second condition was that both model children achieved the same low score of 2 out of 10. The third condition was that the continuous worker model received a 2 while the other received a 10 out of 10. Each subject was then asked a series of four questions to ascertain attributions of the results to cleverness or hard work. Results showed that children move through four phases of attribution. The first phase is an undifferentiated view of effort, ability, and outcome. That is, they do not distinguish between effort and ability for achievement outcome. In the second phase, outcome is attributed to effort, but ability is not yet identifiable. Ability distinction begins in the third phase and, finally, the distinction between effort and ability for outcome becomes clear in the fourth phase.

Attribution of effort or ability for achievement relates to implicit theory of intelligence, and whether a person believed ability was fixed (entity) or malleable (incremental), and able to be improved (Dweck & Leggett, 1988). Implicit theories correspond with distinct response patterns to challenges. These patterns are categorized as maladaptive (helpless) or adaptive (mastery-oriented) responses to failure. Maladaptive response patterns in affect, cognition, and behavior inhibit effective functioning in the face of challenge, such as giving up, or attributing failure uncontrollable causes, such as lack of ability, task difficulty, or luck. Adaptive patterns involve seeking out challenges, generating and utilizing effective strategies to overcome them, and attributing failure to lack of effort (Dweck, 1975). Therefore, one's implicit theory would determine attributions for success or failure on tasks. Given a fixed view,

achievement success or failures would be attributed to ability, while a malleable view would attribute achievement outcomes to effort.

**Response patterns.** Elliott and Dweck (1988) set out to test whether different types of goals set up the helpless or mastery-oriented response patterns using 101 fifth-grade children (57 girls and 44 boys). Researchers first manipulated ability perceptions by giving predetermined feedback after an initial pattern recognition task. The participants were told that they currently had either high or low ability on the task, which divided the group into high ability and low ability perceptions. Next, subjects were asked to choose between a learning task and a performance task. The researcher offered the following distinction in the instructions to the children:

Performance task: In this box we have problems of different levels. Some are hard, some are easier. If you pick this box, although you won't learn new things, it will really show me what kids can do.

Learning task: If you pick the task in this box, you'll probably learn a lot of new things. But you'll probably make a bunch of mistakes, get a little confused, maybe feel a little dumb at times – but eventually you'll learn useful things. (Elliott & Dweck, 1988)

Additionally, researchers informed the children that those who chose the performance task would be filmed, while no filming was mentioned for the learning task. It was assumed that filming would encourage displaying competence, rather than developing competence. The performance task required a further decision to choose moderately easy, moderate, and moderately difficult levels. Despite the distinction in the instructions, the

same pattern discrimination task was used regardless of what the subject selected. Task choice (learning or performance), problem solving effectiveness (based on researcher's questions) and spontaneous verbalizations during the discriminating task, were the dependent variables. Results from the four conditions, high perceived ability-learning task, low perceived ability-learning task, high perceived ability-performance task, low perceived ability-performance task, supported the researchers' hypothesis that different goals set up different response patterns. Under the learning goals condition, neither high nor low perceived ability subjects were likely to make statements of failure attributions. In contrast, of the subjects in the performance condition, 27% of the low perceived ability made attribution statements of failure compared to only 4% of the high-perceived ability. These attributions included lack of ability, lack of effort, task difficulty, experimenter's unfairness, or lack of luck. Interestingly, the attribution statements made by subjects in the low performance-low perceived ability condition were all referenced to uncontrollable causes, rather than effort, which is controllable. The researchers concluded that individuals, who exhibit this pattern of attribution to uncontrollable causes, or helplessness, might pursue different goals than those who exhibit the mastery-oriented patterns. As such, these goals are categorized into the dichotomous performance and learning goals.

**Evolution of the theory.** The dichotomous view of achievement goals has since evolved over time. The original view relied solely on the distinction of task- or norm-referenced definition of competence. Definitions of competence relative to task or one's improvement are mastery goals, while those relative to comparing one's ability to others

are performance goals. Subsequent research split the performance goals into approach and avoid motivations (Cury, 2004; Daniels et al., 2008; Elliot, 1999; Pintrich, 2000). Approach and avoid motivations are distinct functions of valence, and whether behavior is directed toward positive or negative possibilities, respectively (Elliot, 1999). Thus, according to the trichotomous model, the three goals are mastery, performance approach, and performance avoid. Mastery goals continue to be identified with beneficial processes and response patterns surrounding the desire to improve ability. Performance approach goals continue to represent the performance goals already discussed, namely the aim to demonstrate ability relative to others. Performance avoid goals, then, are aimed at avoiding the demonstration of incompetence relative to others.

In a French physical education study, Cury (2004) tested the trichotomous model by using data gathered from questionnaires to correlate goal orientation, perceptions of physical education competence, implicit theories about ability, and perceived motivational climate. Subjects were 682 boys from five French high schools with an age range of 13-16 years. Three distinct patterns of perceptions of physical education competence, entity or incremental beliefs about ability, and perceived motivational climate emerged, and supported the three goal orientations.

Elliot and McGregor (2001b) set out to extend the trichotomous model by applying valence to mastery goals. By examining antecedents and consequences, mastery avoid was investigated as distinct from mastery approach based once again on the dimension of valence. The researchers conceptualized the two fundamental dimensions of competence as definition and valence. Definition of competence is reference to an



absolute standard (task), an intrapersonal standard (self), or a normative standard (others). Based on prior literature the researchers consider absolute and intrapersonal standards together, and continue the mastery and performance components of the original achievement goal dichotomy. However, while the trichotomous model only applied the valence component to performance goals, Elliot and McGregor further applied valence to mastery goals, which gave rise to the mastery avoid goals. Mastery-avoid goals share the same competence definition as mastery approach goals, which is task/self referenced. The difference lie in the valence, with approach oriented to positive consequences, such as improvement and mastery; and avoid oriented to evade negative consequences, such as errors, forgetting, losing capabilities, or missing basketball free throws. As a result, the 2 x 2 achievement goal framework was proposed and tested. Subjects were 49 male and 131 females undergraduates who were given extra credit in an introductory psychology class. The 2 x 2 achievement goal questionnaire was devised as a 12-item instrument with each question based on a 1-7 scale indicating how true the statements were for the participant. Three questions for each dimension, including the new mastery avoid, were chosen based on previous pilot studies. Exploratory factor analysis showed that mastery avoid was operative but less prevalent than the other three forms of goals. Additional studies included in the research paper reported that mastery avoid showed correlations with goals that shared one of the dimensions of definition or valence (mastery approach or performance avoid), but not with the unrelated performance approach. Further, mastery avoid was associated with more negative process than mastery approach, but more positive than performance avoid. Given the complexities of competitive sports, it is

worth continuing to investigate the 2 x 2 framework and the mastery-avoid goal orientation.

As achievement goal theory evolved, the dichotomous view was also revised with the possibility that a person can hold multiple goals. This orthogonal view was examined by Pintrich (2000) in a math study with 8<sup>th</sup> and 9<sup>th</sup> graders. Only mastery and performance-approach goals were considered in the study. The researcher administered a self-report questionnaire in three waves, occurring October and May of 8<sup>th</sup> grade and then again in May of 9<sup>th</sup> grade. The questionnaires included subscales of mastery beliefs, affect, and strategy use. Goal orientation was also measured during the first wave of the questionnaire administration. A median split procedure was used to divide the mastery and performance-approach results into high and low categories. Thus, the participants were divided into four groups: low-mastery/low performance, low-mastery/high-performance, high-mastery/low-performance, and high-mastery/high-performance. It was shown that the high-mastery/high-performance group had similar, if not better, outcomes in terms of motivational beliefs, affect, strategy use, and performance than the high-mastery/low-performance group. The finding supported the revision of the dichotomous view into a multiple goals theory of goal orientation.

**Achievement goals in sport.** Achievement goals theory studies have gained some traction in the sport domain. Cross-domain associations among personal goals and beliefs about the causes of success, perceived ability, satisfaction, and boredom were examined between classroom and sport in 207 (99 male, 108 female), 10<sup>th</sup> and 11<sup>th</sup> grade high school students. In both domains, goal orientation correlated with expected beliefs for

success (Duda & Nicholls, 1992). That is, performance oriented individuals attributed success to ability, while mastery oriented individuals attributed success to effort, interest, and collaboration in both classroom and sport. Results obtained in a sport specific study by Cervelló and Santos-Rosa (2001) with 300, Spanish 15-17 year-olds from a variety of recreational-level sports also showed that high ego (performance) and low task (mastery) orientation were associated with lower perceived ability, preference for easy tasks, and less enjoyment and satisfaction.

More recent sport studies on AGT tend to focus more on the perceived motivational climate of the group along with specific coaching feedback behaviors associated with perceived mastery or performance climates (Seifriz, Duda, & Chi, 1992; A. L. Smith, Balaguer, & Duda, 2006; R. E. Smith, Smoll, & Cumming, 2009), which is outside of the intent of the current study. In addition, a majority of the existing studies are based on youth or adolescent participants. Few involve collegiate student-athletes, and these are concerned with Ryan and Deci's (2000) self-determination theory (Amorose & Anderson-Butcher, 2007; Amorose & Horn, 2000, 2001), or with recreational, rather than competitive, sport activity (Conroy, Elliot, & Hofer, 2003). The distinction between age and level of collegiate athletes would be important based on the impressionistic beliefs of the younger groups.

The review of literature to this point is devoid of any explicit connections between calibration and AGT. However, the relationships between the two are prevalent, though implicit. As evidence, I offer passages from articles from two revered researchers on AGT. The first is from C. Ames (1984):

Because they were expecting another task and because ability is generally perceived as stable and predictive of one's future performances, subjects who attributed their performance to ability should expect to perform at a similarly high or low level on the upcoming task. (p. 485)

On what basis is ability “predictive” of “future performances”, and how do we know if subjects performed as expected? The answer is found in calibration research. The second passage is from Nicholls (1984):

To ascertain the most economical action for demonstrating ability, individuals must form subjective probabilities of demonstrating high versus low ability on the available tasks. (p. 332)

To be clear, the difference between “subjective probabilities” and “demonstrating...ability” is calibration.

In a recent study on metacomprehension and self-regulation among undergraduates, Zhou (2013) explored the proposed connections between calibration and goal orientation in reading comprehension tasks. Elliot and McGregor's (2001a) Achievement Goal Questionnaire was used to measure goal orientation. Reading comprehension calibration ratings were collected at three different times (after sample, after initial reading, after review) and compared with actual performance on 10-item test following the review reading. Cluster analysis was used to identify achievement goal and calibration confidence profiles among the participants. High-monitoring-overconfidence, low-monitoring-overconfidence, and moderate-monitoring-underconfidence groups emerged, supporting a multiple-goals perspective. However, calibration accuracy was not

differentiated among goal profiles. Performance-approach goals were found to positively relate to overconfidence. The authors suggested that students with performance approach goals were more interested in projecting a smart image, and therefore more likely to express higher confidence in their ability to perform on a test.

Now that the relationship between calibration and AGT is established, we turn to the issue of generalizability of findings to other domains. Undergraduate students, similar in some respects to collegiate student-athletes, dominate the research in calibration. However, as previously discussed, the two populations differ in many important ways, including, but not limited to, the goals of competitive student-athletes to concurrently develop and demonstrate ability, and the normative nature of this group's perception of ability.

## **Purpose**

The purpose of this study was to examine achievement goal orientation patterns and relationships with calibration, expertise, and performance in competitive, collegiate volleyball players at NCAA member universities.

The 2 x 2 achievement goal orientation framework provides the most opportunity for categorizing different goals. Therefore, when combined with the possibility that a person can pursue multiple goals (Pintrich, 2000) rather than just one type, one of the aims of the study was to examine potential achievement goal patterns. Of particular interest were the patterns that may exist relative to experience and to performance. Given that the population for the current study is inherently more concerned with performance goals, it was expected that the subjects in the study would show overall high levels of performance goal orientations.

The comparison of prediction to achievement should occur prior to attribution of effort or ability for the achievement. Additionally, striving for improving or proving ability implies an initial state of accurately estimating current ability. Players who are working for improvement can only judge the extent of improvement relative to a baseline of ability. Likewise, players who are working to prove their ability must know their current ability in order to judge whether or not it has been reached, or proven. Considering the variety of the potential relationships among calibration accuracy,

calibration confidence, experience, performance, and goal orientation, expectations for specific results were not made.

Despite the methodological and conceptual difficulties, calibration research has consistently shown a pattern of overconfidence with respect to accuracy. Another pattern to emerge is the hard/easy effect, which is the tendency for people to be overconfident for hard tasks, and underconfident for easy tasks. That is, people overestimate their performance for those tasks that are difficult, and underestimate for those that are easy. It was expected that both the overconfidence pattern and the hard/easy effect would generally be seen in competitive volleyball players. Additionally, the differences in these patterns relative to performance were explored.

Based on the expected relationships, the following exploratory questions were examined:

R1. What patterns of goal orientation exist?

- a. How did these patterns relate to experience?
- b. How did these patterns relate to performance?

R2. What were the relationships among calibration accuracy, calibration confidence, experience, performance, and goal orientation?

R3. What patterns of calibration exist?

- a. How did these patterns relate to performance?

## **Methods**

### **Participants**

The participants were 72 male volleyball players (ages 18-23 years, mean = 19.78, SD = 1.37) from seven different college institutions from within and neighboring the Mid-Atlantic region of the United States. The frequencies, means, and standard deviations of the demographic information are displayed in Table 1. Most of the participants were White/Caucasian (84.7%). All of the participant programs were Division 1, 2, and 3 (30.6%, 25.0%, 44.4%, respectively) members of the National Collegiate Athletic Association (NCAA) and conveniently sampled as a result of their participation in a tournament held at a university in Fairfax, Virginia on November 23, 2013. The number of players who volunteered from each program ranged from six to twenty. Two of the programs split their roster to create two teams each in order to maximize the number of competitive opportunities during the tournament. Various experience levels were represented, as the sample ranged from first to fifth year college volleyball players with various previous experiences. While the host institution participated in the tournament, these players did not participate in the study due to the researcher's position as their coach.

### **Measures**



A questionnaire devised specifically for the study was used. Demographic questions were typical of such data collection. An Achievement Goal Orientation – Revised questionnaire (AGQ-R), modified for the volleyball domain, was used in addition to a series of vignettes to determine goal orientation (see Appendix A for Questionnaires). Performance was assessed using data collected from each match played.

**Demographic information.** Information of age, ethnicity, academic class, and NCAA division designation of the volleyball program was collected from the players in the first four questions of the questionnaire. The demographics of the coaches were not required as they served as gatekeepers and data collection facilitators.

**Experience.** An experience index was calculated from questionnaire items that asked for the number of years playing various levels of volleyball. Each category (high school, club, USA High Performance, and college) was weighted from 1 (high school) to 4 (college), multiplied by the number of years indicated, and summed to yield an experience index.

**Achievement goal orientation.** A modified version of the Achievement Goal Questionnaire-Revised (AGQ-R) (Elliot & Murayama, 2008) was used to examine individual's achievement goal orientation. Responses to items such as “My goal is to perform better than the other players” were indicated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The modifications incorporate changes in verbiage to adapt original questions to the sport domain, similar to the sport and physical education modifications (Conroy et al., 2003; Wang, Biddle, & Elliot, 2007) made to the original Achievement Goal Questionnaire (Elliot & McGregor, 2001b). The AGQ-R was

shown to be a better fit for the competence and valence model of achievement goals than the original Achievement Goal Questionnaire (Elliot & Murayama, 2008). The 12-item questionnaire was divided evenly to measure the four types of goals. Cronbach's alphas for the mastery approach, mastery avoid, performance approach, and performance avoid subscales were .68, .86, .61, and .93, respectively. Evidence for validity has been previously demonstrated (Elliot & Murayama, 2008).

Vignettes were created as an additional measure of achievement goal orientation. Each vignette was constructed with statements that reflected self- or norm-based comparisons of competence, positive or negative valence, and ability or effort attributes. Players indicated their level of agreement with each vignette on a 1-5 Likert scale. The following is a vignette constructed as an indication of performance approach level:

I like knowing how I played compared to the other players. I work hard because I want to be the best on my team. I practice harder when my coach posts everyone's performance. My performance has more to do with my ability than my effort. If I play in a match, I look at match statistics to see how I did compared to everyone else. If I don't perform well, it's usually because I was unlucky.

Circle your answer

|                    |   |   |   |               |
|--------------------|---|---|---|---------------|
| 1                  | 2 | 3 | 4 | 5             |
| not at all like me |   |   |   | a lot like me |

Between the questionnaire items and the vignettes, mastery-approach, performance-approach, mastery-avoid, and performance-avoid showed mild to moderate, positive correlations ( $r(69) = .24$ ,  $r(70) = .30$ ,  $r(70) = .30$ , and  $r(70) = .43$ , respectively,  $p < .05$ ).

**Perceived opponent difficulty.** Two items on the questionnaire were used to measure a perceived level of opponent difficulty. The first was quantitative based and

participants were asked, “How would you rate the ability of your upcoming opponent?” Answers were based on a 5-point Likert scale ranging from 1 (low ability) to 5 (high ability). The second item was qualitative in nature and was an open-ended question that asked, “What is this rating based on?” The rationale for the measure was to obtain a measure of perceived difficulty as seen in calibration studies based on questions of knowledge.

**Performance.** Coaches or staff members of each team collected attack performance data during each of the three matches. An attack is an offensive ball contact whereby a player jumps to strike the ball at maximum height with the intent to score. Collected attack data were typical of information collected during a match for team analysis. Therefore, collection training was not necessary. A Match Data Sheet was collected from each team for the three matches, corresponding to the three questionnaires that were administered (see Appendix B for Match Data Sheet). Specifically, the actual number of “Kill,” “In Play,” “Blocked,” and “Error” for each player was kept. A “kill” is an attack that directly results in a point scored for the attacking team. This occurs if the attacked ball hits the opponent’s court or contacts an opponent in a way that prohibits further controlled contact by other members of the opponent team. An “in play” refers to an attacked ball that is controlled by the opponent. A “blocked” attack is one that is stopped by the opponent’s block and cannot be recovered. Finally, an “error” is an attack that lands out of bounds, fails to clear the net, or contacts the antennae (out of bounds marker at the net). In the NCAA, it is common for blocked and errors to collapse into a general “error”. While statistically treated as the same in volleyball, they are conceptually

different in that blocked attacks are “forced” errors by the opponent, while attack errors are “unforced”. However, the conceptual difference still yields a point for the opponent, which is of primary concern to the study. Therefore, in order to use the standard calculation, “blocked” and “errors” were added together, and the following equation was used to measure actual performance:

$$efficiency_{actual} = \frac{(kills - errors)}{total\ attempts}.$$

The possible range of efficiency using this calculation is -1.00 (worst possible) to +1.00 (best possible).

**Calibration accuracy.** On each of the questionnaires, players were asked to estimate their individual performance on the skill of attacking. Players indicated their personal estimates in percentages for each of the possible attack outcomes by filling in the question:

What do you estimate will be your percentages among the following categories for YOUR ATTACKS?

|                               |                   |                   |                   |                   |                               |
|-------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------------------|
|                               | <u>        </u> % | <u>        </u> % | <u>        </u> % | <u>        </u> % | <u>        </u> %             |
| (total should add up to 100%) | Kill for point    | In play           | Blocked           | Error             | I do not attack (mark with X) |

Estimated percentages of attack Kills, In Play, Blocked, and Errors were converted to an attack efficiency percentage, again using the volleyball standard calculation:

$$efficiency_{estimated} = \frac{(kills - errors)}{total\ attempts}.$$

Because the estimates are given in the form of percentages, the total attempts denominator is 100. Performance estimates were collected prior to a match on three separate occasions throughout the day. The discrepancy score was calculated for each of the three matches, and the mean was calculated to obtain an average discrepancy measure for calibration accuracy.

Calibration accuracy was then calculated as the absolute difference between estimated and actual performance. Similar to the terminology used by Wahlstrom (2001), “discrepancy” was used as an index of accuracy, with high discrepancy values corresponding to low accuracy. Thus, the following equation was used for indicating calibration accuracy:

$$discrepancy = |efficiency_{estimated} - efficiency_{actual}|.$$

The potential range of discrepancy using this calculation is 0 to 2.00, with 0 representing no discrepancy (perfect accuracy) and 2.00 representing extreme discrepancy (no accuracy). For example, a player who estimates a best possible performance of 1.00 but actually obtains a worst possible performance of -1.00 would have a discrepancy score of 2.00 and demonstrate no accuracy. Because of the absolute value, this result would be the same if the estimates and actuals were at the opposite extremes.

**Calibration bias.** Calibration bias is simply the difference between performance estimates and performance actuals, represented by the equation:

$$calibration\ bias = efficiency_{estimated} - efficiency_{actual}$$

The potential range of bias is -2.00 to 2.00. Maintaining the signed difference differentiates bias from accuracy (discrepancy), with positive numbers indicating over-calibration (overconfidence) and negative numbers indicating under-calibration (underconfidence). Calibration bias was calculated for each player for each match, and then averaged for an average bias index.

**Calibration confidence.** After each group of estimations on the questionnaires, participants were asked to give a “% confidence” rating for their estimates. The mean confidence ratings were calculated across the three matches.

## **Procedures**

IRB approval was obtained before conducting the study. At an initial meeting with only the coaches, the researcher introduced the study and explained the voluntary nature of participation. The coaches were asked for their participation first, as they would be an integral component to the data collection. It was further explained that players of coaches who did not agree to participate would not be asked to participate. All seven coaches agreed to participate. Written consent was then collected from the coaches (see Appendix C for Informed Consent). Coaches then completed the first step of the data collection, which was to assign their players to a Player Identification Key. Coaches were instructed to use the key to maintain confidentiality of their players on all of the collected data. Instructions for collecting and reporting match data for three matches were given. Blank player questionnaires for Match 2 and Match 3 were then given to the coaches for each of their players with instructions to have their volunteers fill them out prior to

upcoming matches throughout the day. The first questionnaire for the first match was to be introduced and collected directly by the researcher.

Next, a meeting was held with the players from all of the teams to introduce and explain the study. It was once again presented as voluntary, and there were no benefits for participation and no consequences for non-participation. The process for maintaining their confidentiality with Team and Player ID's was explained, and players were then given the informed consent and the first questionnaire. Those who agreed to participate were asked to read and sign the informed consent and return these to the researcher. Finally, the volunteers were asked to fill out the questionnaire for Match 1. Upon collection of the questionnaires, the meeting was completed and the players were released to begin the tournament. Because the first questionnaire referred to estimates for an upcoming match, the first match of the day for each team thus became the first opportunity to collect actual match data.

The questionnaires for Match 2 and Match 3 were disseminated to the players prior to their respective matches by their coach rather than the researcher, as the nature of the tournament format prohibited a large group interruption in the schedule. The researcher and a volunteer assistant circulated throughout the tournament and met informally with the coaches to collect the questionnaires throughout the day. Finally, the actual attack results from the three matches were collected from the coaches at the end of the tournament.

## **Results**

Analysis of the data began with examination of descriptive statistics of the measures. Patterns of goal orientation were examined using cluster analysis. Independent-samples t test was used to examine differences in clusters with respect to experience and performance. Correlation was used to examine relationships among calibration accuracy, confidence in estimates of performance, actual performance, and goal orientations. Patterns of calibration were explored with graphical analysis.

All volunteers who completed the first questionnaire and provided data to be examined for achievement goal orientations were counted toward the initial sample size of 72. However only data from players who provided performance estimates ( $n = 51$ ) and performed the skill of attacking ( $n = 52$ ) in the recorded matches were examined for calibration accuracy. Of the initial 72 participants, 57 provided confidence ratings for estimations. Due to either missing or incomplete data, the sample size was reduced 49 when examining patterns of calibration.

### **Descriptives**

Demographic information is displayed in Table 1. The demographic characteristics of the sample appear typical of the NCAA men's volleyball population. Variations in the total number of responses for each characteristic were due to missing responses for items on the questionnaire.



Table 1

*Demographic Information*

| Characteristic                            | Frequency | Percent | Mean  | SD   |
|---|-----------|---------|-------|------|
| Age                                       |           |         | 19.78 | 1.36 |
| 18.0                                      | 19        | 26.4    |       |      |
| 19.0                                      | 10        | 13.9    |       |      |
| 20.0                                      | 19        | 26.4    |       |      |
| 21.0                                      | 17        | 23.6    |       |      |
| 22.0                                      | 6         | 8.3     |       |      |
| 23.0                                      | 1         | 1.4     |       |      |
| Total                                     | 72        | 100.0   |       |      |
| Ethnicity                                 |           |         |       |      |
| Black/African American                    | 3         | 4.2     |       |      |
| White/Caucasian                           | 61        | 84.7    |       |      |
| Asian                                     | 2         | 2.8     |       |      |
| Other                                     | 4         | 5.6     |       |      |
| Mixed ethnicity                           | 2         | 2.8     |       |      |
| Total                                     | 72        | 100.0   |       |      |
| Class                                     |           |         |       |      |
| Freshman                                  | 20        | 27.8    |       |      |
| Sophomore                                 | 15        | 20.8    |       |      |
| Junior                                    | 22        | 30.6    |       |      |
| Senior                                    | 14        | 19.4    |       |      |
| Fifth year                                | 1         | 1.4     |       |      |
| Total                                     | 72        | 100.0   |       |      |
| NCAA division                             |           |         |       |      |
| Division 1                                | 22        | 30.6    |       |      |
| Division 2                                | 18        | 25.0    |       |      |
| Division 3                                | 32        | 44.4    |       |      |
| Total                                     | 72        | 100.0   |       |      |
| Years high school experience <sup>a</sup> |           |         | 2.80  | 1.84 |
| 0.0                                       | 17        | 23.6    |       |      |
| 1.0                                       | 2         | 2.8     |       |      |
| 2.0                                       | 4         | 5.6     |       |      |

|                                       |    |      |      |      |
|---------------------------------------|----|------|------|------|
| 3.0                                   | 10 | 13.9 |      |      |
| 4.0                                   | 33 | 45.8 |      |      |
| 5.0                                   | 1  | 1.4  |      |      |
| 6.0                                   | 2  | 2.8  |      |      |
| 7.0                                   | 1  | 1.4  |      |      |
| Total                                 | 70 | 97.2 |      |      |
| Years club experience                 |    |      | 3.30 | 2.59 |
| 0.0                                   | 10 | 13.9 |      |      |
| 1.0                                   | 9  | 12.5 |      |      |
| 2.0                                   | 12 | 16.7 |      |      |
| 3.0                                   | 9  | 12.5 |      |      |
| 4.0                                   | 10 | 13.9 |      |      |
| 5.0                                   | 6  | 8.3  |      |      |
| 6.0                                   | 7  | 9.7  |      |      |
| 7.0                                   | 3  | 4.2  |      |      |
| 8.0                                   | 2  | 2.8  |      |      |
| 9.0                                   | 1  | 1.4  |      |      |
| 13.0                                  | 1  | 1.4  |      |      |
| Total                                 | 70 | 97.2 |      |      |
| Years USA high performance experience |    |      | 0.68 | 1.04 |
| 0.0                                   | 43 | 59.7 |      |      |
| 1.0                                   | 12 | 16.7 |      |      |
| 2.0                                   | 8  | 11.1 |      |      |
| 3.0                                   | 5  | 6.9  |      |      |
| 4.0                                   | 1  | 1.4  |      |      |
| Total                                 | 69 | 95.8 |      |      |
| Years college experience              |    |      | 1.64 | 1.27 |
| 0.0                                   | 16 | 22.2 |      |      |
| 1.0                                   | 18 | 25.0 |      |      |
| 2.0                                   | 18 | 25.0 |      |      |
| 3.0                                   | 11 | 15.3 |      |      |
| 4.0                                   | 7  | 9.7  |      |      |
| Total                                 | 70 | 97.2 |      |      |

*Note.* Variations in totals are a result of missing answers to questionnaire items.

<sup>a</sup>The three players who have more than four years of high school volleyball experience may have played at high school level while in middle school.

As seen in Table 1, experience varied across levels. For high school volleyball, almost half indicated four years of experience (45.8%), with the next closest group with zero years of high school volleyball (23.6%). This is not unexpected, as there are few high schools that sponsor boys' volleyball. Club level showed the highest percentage of participants with 2 years of club (16.7%) followed by 0 and 4 years (13.9%). Over half of the participants had no experience in the USA High Performance program (59.7%), with the next largest group with only one year (16.7%). Experience at the college level was more evenly distributed, with the highest frequencies in 0, 1, and 2 years of collegiate experience (22.2, 25.0, and 25.0, respectively).

A summary of the descriptive statistics for the measures is displayed in Table 2.

Table 2

*Summary of Measures Descriptives*

| Measure                                | n  | Min  | Max   | Mean  | SD    |
|--|----|------|-------|-------|-------|
| Experience index <sup>a</sup>          | 70 | 1.00 | 39.00 | 17.99 | 10.12 |
| Mastery approach questionnaire         | 72 | 3.00 | 5.00  | 4.82  | .36   |
| Mastery avoid questionnaire            | 72 | 1.00 | 5.00  | 4.13  | 1.15  |
| Performance approach questionnaire     | 72 | 3.00 | 5.00  | 4.59  | .52   |
| Performance avoid questionnaire        | 72 | 1.00 | 5.00  | 3.82  | 1.27  |
| Mastery approach vignette <sup>b</sup> | 71 | 1.00 | 5.00  | 4.07  | .92   |
| Mastery avoid vignette                 | 72 | 1.00 | 5.00  | 3.54  | 1.20  |
| Performance approach vignette          | 72 | 1.00 | 5.00  | 3.46  | 1.21  |
| Performance avoid vignette             | 72 | 1.00 | 5.00  | 3.21  | 1.20  |
| Perceived opponent difficulty 1        | 71 | 1.00 | 5.00  | 3.83  | .86   |
| Perceived opponent difficulty 2        | 49 | 1.00 | 5.00  | 3.65  | 1.05  |
| Perceived opponent difficulty 3        | 48 | 1.00 | 5.00  | 3.37  | .87   |
| Performance estimate 1                 | 55 | -.80 | .70   | .19   | .26   |

|                              |    |       |        |       |       |
|------------------------------|----|-------|--------|-------|-------|
| Performance estimate 2       | 45 | -.71  | .80    | .18   | .31   |
| Performance estimate 3       | 45 | -.30  | .75    | .27   | .26   |
| Average performance estimate | 58 | -.76  | .70    | .20   | .26   |
| Actual performance 1         | 50 | -.75  | 1.00   | .18   | .47   |
| Actual performance 2         | 49 | -1.00 | 1.00   | .21   | .37   |
| Actual performance 3         | 48 | -1.00 | 1.00   | .17   | .38   |
| Average performance          | 52 | -.58  | .88    | .18   | .30   |
| Discrepancy 1                | 46 | .02   | 1.20   | .42   | .31   |
| Discrepancy 2                | 40 | .00   | 1.40   | .35   | .28   |
| Discrepancy 3                | 40 | .00   | 1.20   | .29   | .25   |
| Average discrepancy          | 51 | .08   | 1.40   | .36   | .26   |
| Bias 1                       | 46 | -1.20 | 1.03   | -.01  | .53   |
| Bias 2                       | 40 | -1.40 | 1.03   | -.02  | .45   |
| Bias 3                       | 40 | -.70  | 1.20   | .10   | .37   |
| Average Bias                 | 51 | -1.40 | .86    | .00   | .39   |
| Confidence estimate 1        | 54 | 10.00 | 100.00 | 77.83 | 22.84 |
| Confidence estimate 2        | 44 | 20.00 | 100.00 | 76.11 | 20.64 |
| Confidence estimate 3        | 42 | 20.00 | 100.00 | 79.26 | 20.17 |
| Average confidence           | 57 | 18.33 | 100.00 | 78.32 | 19.37 |

*Note.* Variations in sample size are due to the variation in the number of subjects who completed subsequent questionnaires, and who actually performed the attacking skill.

<sup>a</sup>Experience was calculated for each player as the weighted sum of high school, club, USA high performance, and college experience. Weighted multipliers were 1, 2, 3, and 4, respectively, representing higher-level experience with higher multipliers. <sup>b</sup>One of the subjects failed to indicate a level for the item.

As shown in Table 2, results from the modified AGQ-R indicate highest levels of the two approach orientations, with mean levels for mastery-approach at 4.82 and performance approach at 4.59. Additionally, no participant indicated lower than a 3 on the 5-point Likert scale for these two measures. Average discrepancy reduced for subsequent matches, while average confidence showed little difference, indicating improved accuracy with similar confidence. This may be indicative of the influence of practice over multiple trials has on improving accuracy.

A correlation matrix for the measures is displayed in Table 3. Discrepancy was negatively correlated with mastery-approach,  $r = -.34$ ,  $n = 51$ ,  $p = .014$ . Confidence was positively correlated with performance-approach,  $r = .29$ ,  $n = 57$ ,  $p = .028$ . The correlations among goal orientations were in the expected directions.

Table 3

*Summary of Correlations*

| Measure                     | 1     | 2    | 3    | 4    | 5     | 6     | 7     | 8     | 9     | 10    | 11    | 12 |
|-----------------------------|-------|------|------|------|-------|-------|-------|-------|-------|-------|-------|----|
| 1. Discrepancy <sup>a</sup> | --    |      |      |      |       |       |       |       |       |       |       |    |
| 2. Confidence <sup>b</sup>  | -.11  | --   |      |      |       |       |       |       |       |       |       |    |
| 3. Experience               | -.06  | .11  | --   |      |       |       |       |       |       |       |       |    |
| 4. Performance <sup>c</sup> | .07   | -.10 | .14  | --   |       |       |       |       |       |       |       |    |
| 5. Mastery Approach Q       | -.23  | .13  | -.21 | .22  | --    |       |       |       |       |       |       |    |
| 6. Mastery Avoid Q          | -.08  | -.06 | -.12 | -.04 | .12   | --    |       |       |       |       |       |    |
| 7. Performance Approach Q   | .00   | -.01 | -.20 | .08  | .48** | .22** | --    |       |       |       |       |    |
| 8. Performance Avoid Q      | .08   | -.02 | -.14 | -.07 | -.00  | .76** | .38** | --    |       |       |       |    |
| 9. Mastery Approach V       | -.34* | -.21 | .18  | .05  | .24*  | .22   | .22   | .17   | --    |       |       |    |
| 10. Mastery Avoid V         | -.17  | -.15 | .10  | -.03 | .04   | .30** | -.08  | .26*  | .35** | --    |       |    |
| 11. Performance Approach V  | -.13  | .29* | .01  | .09  | .06   | .19   | .30*  | .23*  | .34** | .13   | --    |    |
| 12. Performance Avoid V     | -.02  | -.04 | .07  | .03  | -.01  | .38** | .19   | .43** | .27*  | .31** | .53** | -- |

*Note.* A larger sample (N = 72) filled out Questionnaire 1 on calibration, goal orientation, and experience. Items 5-8 obtained from Achievement Goal Questionnaire. Items 9-12 obtained from vignettes. Variations in sample size are due to the variation in the number of subjects who completed subsequent questionnaires, and who actually performed the attacking skill.

<sup>a</sup>n = 51. <sup>b</sup>n = 57. <sup>c</sup>n = 52.

\* p < 0.05. \*\*p < 0.01.

## **Patterns of Goal Orientation (R1)**

Mastery-approach, mastery-avoid, performance-approach, and performance-avoid measures from both the AGQ-R and the vignettes were used in cluster analysis to determine the patterns of goal orientations that exist in the sample. Squared Euclidean Distance was used with Ward's hierarchical cluster analysis to initially examine potential achievement goal profiles from the measures. Examination of the dendrogram to identify the largest distances between clusters suggested two-, three-, and four-cluster solutions. Next, K-means cluster analysis was used for each of the three possibilities. Given the multiple goals possibilities, the differentiation between competence definition and valence within the 2 x 2 achievement goal framework, examination of the agglomeration table, the interpretation of the dendrogram, and the reoccurrence of two clusters, the two-cluster solution was determined to be most appropriate for further analysis.

The first cluster consisted of players who reported high levels of all four-goal orientations and are thus labeled high-all ( $n=52$ ). Cluster two represented players who reported high levels of mastery- and performance-approach and low levels of mastery- and performance-avoid orientations and are labeled high-approach/low-avoid ( $n=19$ ).

Independent-samples  $t$  test was used to compare the clusters of goal profiles in relation to experience and performance. There was no difference in experience between the high-all profile players ( $M = 17.28$ ,  $SD = 10.73$ ) and the high-approach profile players ( $M = 20.11$ ,  $SD = 8.45$ ),  $t(67) = -1.03$ ,  $p = .31$ . Likewise, there was no difference in performance between the high-all players ( $M = .18$ ,  $SD = .31$ ) and the high-approach players ( $M = .18$ ,  $SD = .27$ ),  $t(50) = .28$ ,  $p = .98$ .

## **Relationships Among Calibration Accuracy, Calibration Confidence, Experience, Performance, and Goal Orientation (R2)**

Relationships among accuracy, confidence, experience, performance, and goal orientation were analyzed using correlation. As seen in Table 3, two of the largest correlations were found among items within the Achievement Goal Questionnaire. Mastery-approach was moderately correlated with performance-approach,  $r = .48$ ,  $n = 72$ ,  $p = .000$ , and mastery-avoid was strongly correlated with performance-avoid,  $r = .76$ ,  $n = 72$ ,  $p = .000$ . These results support the 2 x 2 achievement goal framework by the strong relationships between valence (approach and avoid), rather than definitions of competence (mastery and performance). Among the correlations between the questionnaire and the vignettes, correlations exist in the expected directions.

Of the remaining correlations, only two reached statistical significance. From the vignettes, mastery-approach showed a moderate, negative correlation with discrepancy,  $r = -.34$ ,  $n = 51$ ,  $p = .014$ . In other words, higher mastery-approach players were more accurate in their estimates (higher accuracy indexed by lower discrepancy). This result was echoed in the questionnaire, though failed to reach significance,  $r = -.23$ ,  $n = 51$ ,  $p = .112$ . Performance-approach from the vignettes showed a moderate, positive correlation with confidence,  $r = .29$ ,  $n = 57$ ,  $p = .028$ , suggesting that performance-approach players were more confident in their estimates. Both of these results support goal orientation theory. Mastery-approach orientation favors self-referenced definitions of competence, which would yield more accurate estimates of ability. On the other hand, performance-



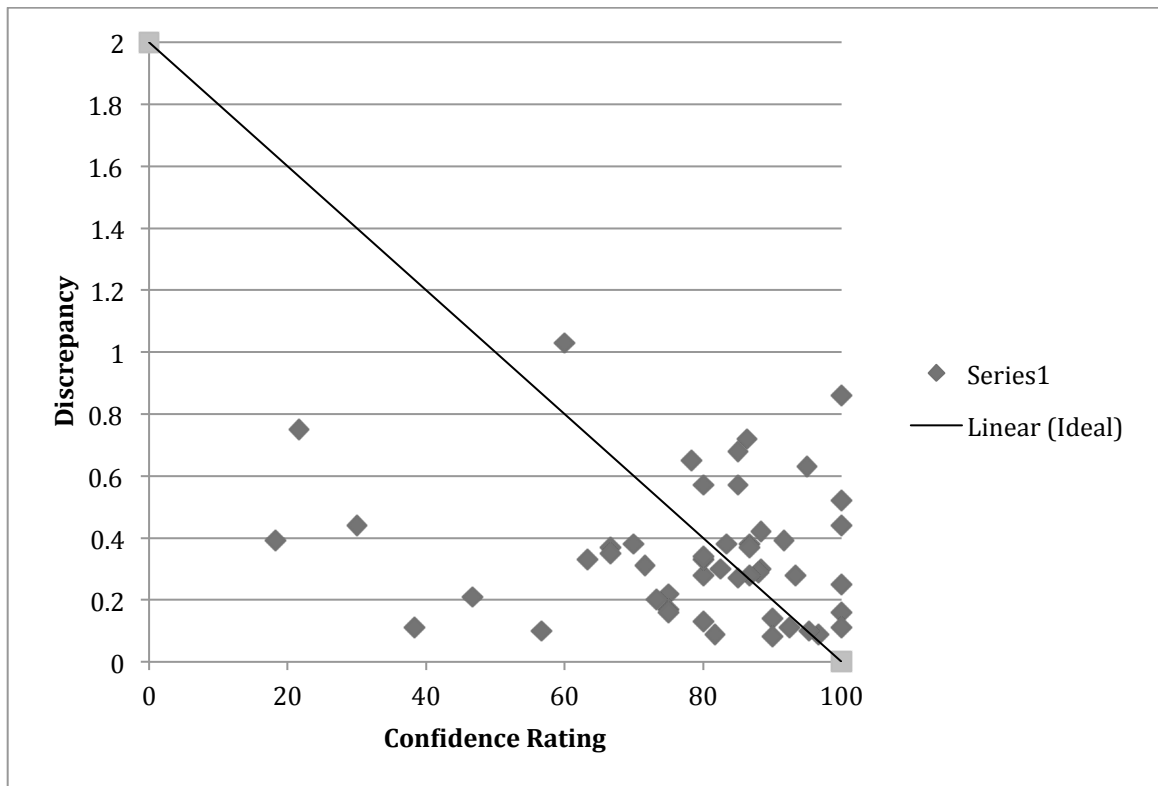
approach defines competence relative to others with a positive valence to appear better than others, which may encourage a higher level of confidence in estimates.

The lack of correlations among discrepancy, confidence in estimates, experience, performance, and goal orientations is an interesting result and will be discussed in a later section.

### **Patterns of Calibration (R3)**

Due to the exploratory nature of the present study, graphical analysis was used to examine the patterns of calibration and expertise. Participants with appropriate data from at least one match were included in the analysis. The sample size for graphical analysis was  $n = 49$ .

**Overconfidence.** In order to examine the patterns of calibration as outlined in the literature, the confidence rating was used similarly to the research in questions of knowledge. Because there were no “correct” or “incorrect” answers in the present study, accuracy was indexed by the inverse of discrepancy. That is, higher discrepancy corresponds to lower accuracy, and vice versa. Figure 1 shows a scatterplot of confidence versus discrepancy for the sample.



*Figure 1.* Scatterplot of confidence versus accuracy (as indexed by discrepancy).

The line in Figure 1 represents the ideal confidence to discrepancy relationship, with 100% confidence corresponding to 0 discrepancy (100% accuracy). Conversely, 0% confidence corresponds to maximum discrepancy index of 2 (complete inaccuracy). Points above this line indicate overconfidence, while points below indicate underconfidence. Visual analysis of Figure 1 shows a high concentration of confidence ratings above the 60% level. Below this level, subjects were all below the ideal calibration line, indicating underconfidence. Above the 60% confidence rating level, both over- and underconfident subjects are seen. In order to clarify the information within Figure 1, the data was rearranged to examine the frequency of subjects within specific

intervals of confidence ratings.

First, 10 intervals of % confidence ratings were created (interval 1 = < 10, interval 2 = 10 to 19.99, interval 3 = 20 to 29.99...interval 10 = 90 to 100). Subjects were placed into these intervals based on their average confidence score. Next, the mean confidence rating and the mean discrepancy for each interval were calculated using each subject's average confidence ratings and average discrepancy score within each interval. A scatterplot of the transformed data is displayed in Figure 2. The numbers next to the points are the number of participants included in the calculation of the discrepancy means.

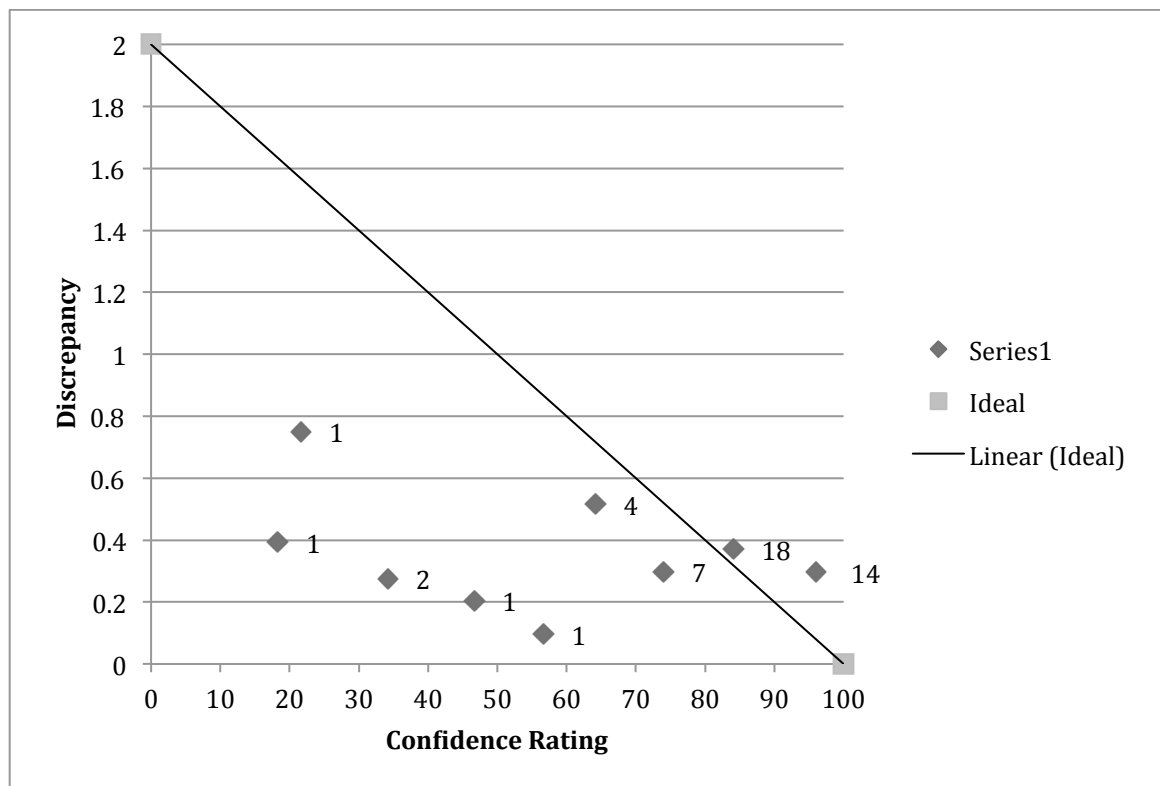
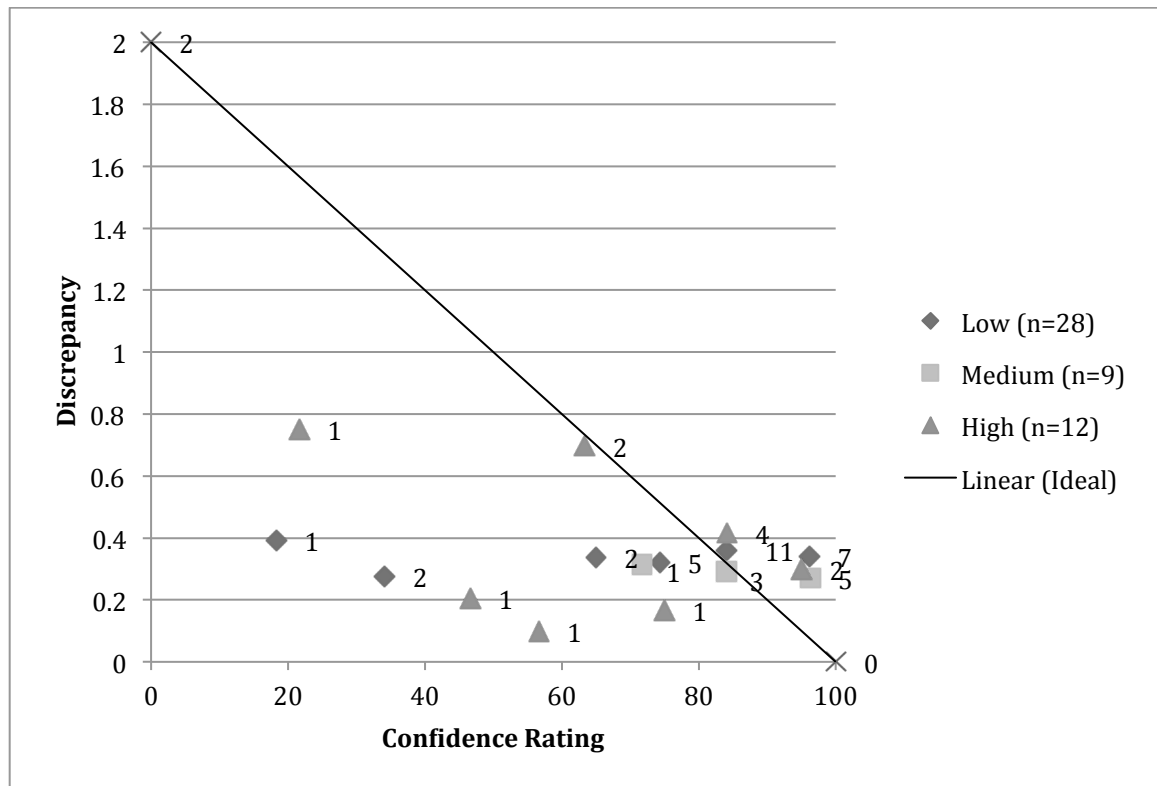


Figure 2. Scatterplot of confidence intervals and discrepancy with frequencies.

As seen in Figure 2, 35% (17) of the players were below the ideal calibration line compared to 54% (32) who were above, when averaged within their respective confidence interval. Thus, Figure 2 suggests that the pattern of overconfidence is evident in the sample, which supports existing research and the initial expectation of the present study. However, within the interval of 80-90% confidence, 37% (18) of the sample appear to be close to the ideal line and could be considered well-calibrated. Consequently, only 29% (14) would be considered overconfident while 35% (17) were underconfident, which would not support the overconfidence pattern.

In order to examine the overconfidence pattern with respect to performance, two graphical analyses was used. The first was an extension of the previous graphical analysis in Figure 2, but the subjects were first grouped into low, medium, and high performing categories. Cut points were created based on the researcher's knowledge of the field and consultation with expert coaches in the field. Low performers had an attack efficiency of less than .25, medium performers from .25 up to .35, and high performers as .35 and above. Similar to Figure 2, the mean discrepancy was plotted against the mean confidence rating for each confidence interval. The difference from Figure 2 lie in that each category of performance was plotted separately in order to examine the calibration patterns with respect to performance. The resultant scatterplot is displayed in Figure 3.

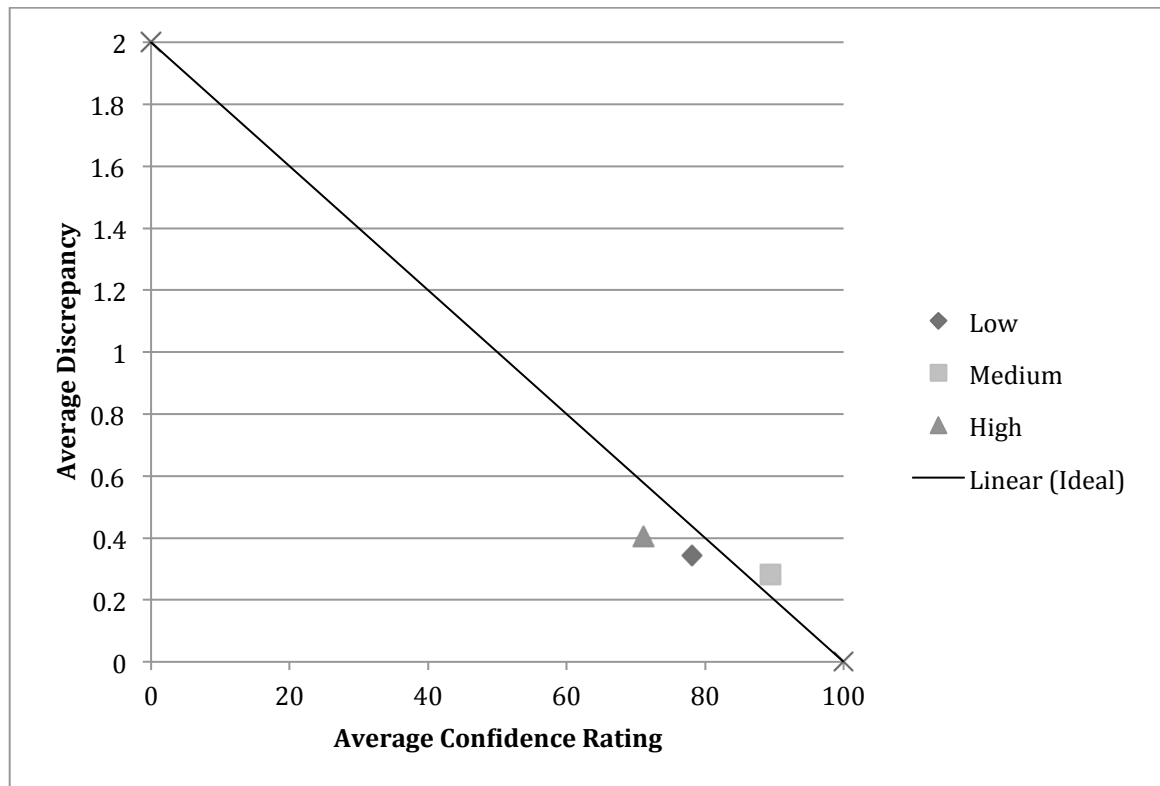


*Figure 3.* Scatterplot of confidence intervals versus discrepancy with respect to performance.

A strict interpretation of Figure 3 shows that on the average for each interval, 36% (10) of the 28 low-performers were underconfident, while 64% (18) were overconfident. Among the medium group, 44% (4) were underconfident while 56% (5) were overconfident. The high-performing group shows 50% (6) as underconfident and 50% (6) as overconfident. The percentages would also support the pattern of general overconfidence for low- and medium-performers. However, as seen in the previous graphical analysis, there are points on the graph that lie very close to the ideal calibration line. If the subjects included in these points are considered well-calibrated, the

breakdown changes. In this case, 36% (10) of the 28 low-performers were underconfident and 25% (7) were overconfident, and 39% (11) were well-calibrated. Among the medium-performers, 11% (1) were underconfident, 56% (5) were overconfident, and 33% (3) were well-calibrated. Finally, among the 12 high-performers, 33% (4) were underconfident, 17% (2) were overconfident, and 50% (6) appear well-calibrated. The results from Figure 3 suggest that the overconfidence pattern only holds for medium performing players, but the low and high performing players were better calibrated with a bias toward underconfidence.

A scatterplot of the relative levels of mean confidence rating and mean discrepancy score for each performance group are displayed in Figure 4.

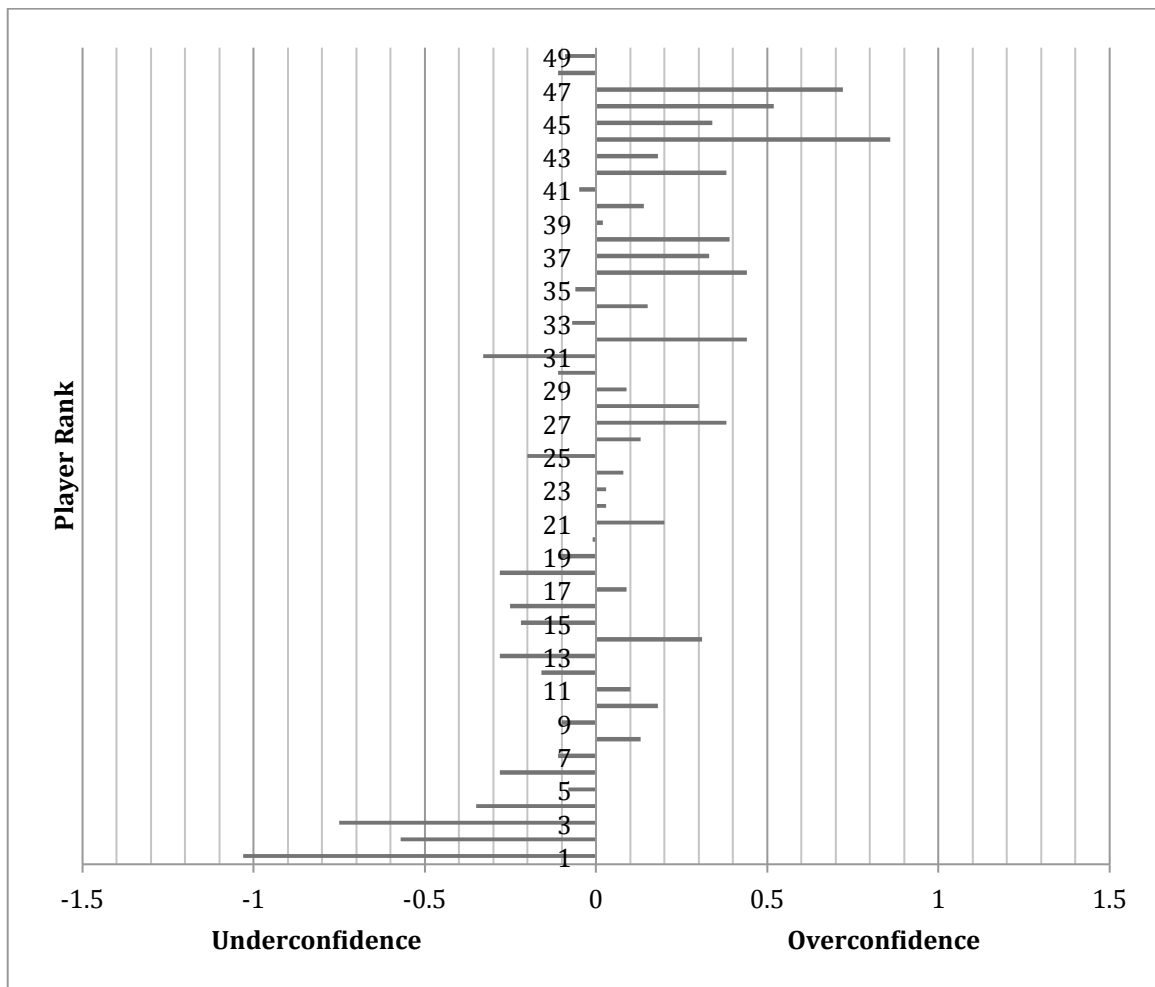


*Figure 4.* Scatterplot of mean confidence versus mean discrepancy for three performance groups.

Figure 4 clarifies the finding that both the low and high performance groups were underconfident, while the medium group was overconfident. In addition, Figure 4 shows that the high-performing group had a lower mean confidence rating, with the low-performing group next lowest, while the medium-performing group gave the highest mean confidence rating. As such, the general pattern of overconfidence does not hold for different levels of performance.

The second graphical analysis used to examine calibration patterns with respect to performance considered the more typical sport calibration model of bias, rather than

confidence in estimates. Players were ranked according to actual performance from 1 (best performer) to 50 (worst performer). As long as subjects played in at least two matches and had corresponding bias scores, they were included for analysis, resulting in a sample size  $n = 49$ . Next, each player's average calibration bias score was plotted on a scale that maintained the signed bias score. Figure 3 represents the average bias of the players ranked by performance.





*Figure 5.* Average bias for each player ranked by performance. Players ranked 1-12 were high-performers, 13-21 were medium-performers, and 22-49 were low-performers.

Figure 5 shows 47% (23) of the 49 subjects were underconfident while 53% (26) were overconfident. While more of the subjects were overconfident, the pattern is not strongly supported.

**Hard-easy effect.** The hard-easy effect was described earlier as the trend in calibration studies to show overconfidence for hard questions and underconfidence for easy questions. Hard items are classified as such because fewer participants answer them correctly, while most participants answer easy items correctly. Translated into sport performance, the hard-easy effect may be described as overconfidence for more difficult tasks (resulting in poorer performance), and underconfidence for easier tasks (resulting in higher performance). Examination of Figure 5 suggests that overconfidence was seen in the poorer performers (more difficult tasks), while underconfidence was seen in the higher performers (easier tasks). This finding offers evidence that supports the initial expectation that the hard-easy effect would be seen in competitive volleyball players.

**Perceived opponent difficulty.** The two questionnaire items used to measure perceived opponent difficulty showed that the quantitative estimates were based on a variety of qualitative factors. Table 4 summarizes the emerging themes based on the responses that were provided from the open-ended question.

Table 4

*Bases of Perceived Opponent Difficulty*

| Theme                           | Sample responses   |
|---------------------------------|--|
| Observation of previous matches | <ul style="list-style-type: none"> <li>• Skill</li> <li>• Previous matches</li> <li>• Their skills I see</li> <li>• Passing, serving, setting, hitting</li> <li>• Film watching and scouting</li> </ul>                              |
| Prior experience                | <ul style="list-style-type: none"> <li>• Previous knowledge of players</li> <li>• Played them last year</li> <li>• Last season's games</li> <li>• Past matchups</li> <li>• Previous competition</li> </ul>                           |
| Program status                  | <ul style="list-style-type: none"> <li>• Collegiate team</li> <li>• College level</li> <li>• They are D1 athletes</li> <li>• Their division</li> <li>• College name and reputation, players, coaches</li> </ul>                      |
| Performance history             | <ul style="list-style-type: none"> <li>• Ability and past performance</li> <li>• Their previous years</li> <li>• Previous season's records</li> <li>• Performance last year</li> <li>• General knowledge, team history</li> </ul>    |
| Appearance                      | <ul style="list-style-type: none"> <li>• Talent, players, size</li> <li>• How they look</li> <li>• Size</li> <li>• First impression, height</li> </ul>   |
| No basis                        | <ul style="list-style-type: none"> <li>• No idea what team I'm playing</li> <li>• I imagine every team is the best, you should never look down on anybody</li> <li>• I don't know</li> <li>• Everyone here is a good team</li> </ul> |

As seen in Table 4, the players based their perceptions of opponent difficulty on external, uncontrollable factors. Of the responses that were given, only one player based his perception with reference his own team by stating, “our ability on our side, not theirs”. A majority of the responses referred to experiences from the previous academic year. Since that time, teams may have lost players to graduation or other reasons. On subsequent questionnaires throughout the day, more of the players based their perceptions on observations of previous matches from the day. However, there was still a portion of responses that were based on history or appearance. Given the nature and variety of responses, further analysis was beyond the scope of the present study.

## Discussion

### Patterns of Goal Orientation (R1)

One of the aims of the present research was to investigate patterns of goal orientation. It was argued that competitive volleyball players represent a motivationally different population than those studied in previous literature. Additionally, due to the performance-dominated nature of competitive athletics, it was expected that performance orientations would prevail. The results did not support this expectation. The results from the cluster analysis suggested a *high-all* group and a *high-approach* group. Both profiles are compatible with the 2 x 2 framework and multiple goals theory. However, there was no difference among the profiles when examined with respect to experience or performance.

Understanding an individual player's profile would provide coaches and sport psychologists with a more specific means of encouraging motivation. The presence of both the high-all and the high-approach profiles suggests that competitive volleyball players have access to various types of motivation, which allows them to maintain high levels of effort, persistence, and use of effective strategies, even when the task changes from mastery (development) to performance (competition). The high-approach profile reflects the division of orientations based on valence, rather than definition of competence. Both of the profiles found in the sample appear typical in sport settings.

Players with the high-all profile would show consistent effort and persistence in both developmental exercises as well as competitive games. These players would be motivated to improve as well as to be better than peers. Additionally, they would be motivated to work for maintenance of ability as well as avoidance of being the worst among peers. From a coach's perspective, these players show consistent improvement and are among the best competitors on the team. They are also the top finishers even when the result is not being recorded. In general, these are the types of players that coaches enjoy working with the most. The high-approach players may fall short of the same levels of motivation for tasks that are more maintenance oriented or directed toward not being the worst. The deficit of a mastery-avoid orientation in this profile would suggest low motivation for tasks designed to avoid losing ability. Likewise, low performance-avoid orientation would suggest low motivation for tasks in which the competitive goal is to meet a minimum criterion, rather than to be the winner. From a coach's perspective, these players would work hard at learning new skills and would "turn on" during competitions in which there is the potential for winning. However, these same players would be more inclined to "go through the motions" of skill fundamental practice and in practices that do not identify a winner. Encouraging motivation for this group would mean clearly defining improvement increments and creating the potential for winning.

### **Relationships Among Accuracy, Confidence, Experience, Goal Orientation, And Performance (R2)**

A second aim of the study was to look at potential relationships among calibration accuracy, confidence, experience, goal orientation, and performance. The complexity of

the potential relationships prohibited stating initial expectations. Out of all of the possible correlations, only two were significant. The first was players with higher mastery-approach levels were more *accurate* in their estimates of performance (as indexed by lower discrepancy levels). A possible explanation is that higher mastery-approach orientation is associated with a clearer sense of current ability, which could lead to more accurate estimates of performance. The second significant correlation was players who were more performance-approach oriented were more *confident* in their estimates of performance. Goal orientation theory informs us that performance-approach players are more inclined to make judgments of ability relative to others. For these players, it may be that because they are accustomed to proving they are better than others, confidence, in general, is more pervasive in their personalities. As a result, they may indicate higher levels of confidence than warranted in all situations, even in the accuracy of their own estimates of performance.

The implications for the relationships between goal orientation, accuracy, and confidence in estimates are also important for player development. The correlation between mastery-approach and accuracy may suggest that these players have a better sense of their current ability, and may be easier for coaches to move through the development process (i.e., more coachable). The second correlation found between performance-approach and confidence in estimates may be indicative of the less coachable player. That is, the high confidence in inaccurate estimates could inhibit a player's ability to train and improve. It is difficult to get players to change what they are doing if they believe they are already doing it correctly. At the highest levels of

performance, this may not be an issue, as improvement is not as imperative as performance. However, at developmental levels, the performance-approach orientation may prevent a player from reaching the highest level of performance.

As noted earlier, the lack of correlations is of particular interest. While the result may be dismissed as simply a lack of correlations, it is more likely that the measures failed to adequately capture the idea of calibration in an appropriate context. The calibration paradigm itself, whereby people answer a question of knowledge and then report a confidence level of correctness, may transfer only to certain sports. For example, a track sprinter may be able to run an event and then, without looking at the recorded time, be able to estimate his or her time and report a confidence in accuracy. The actual result could then be compared to the estimate. The same scenario can be developed for swimmers, or cyclists. The crux of the paradigm is the ability to perform a task, then make an accuracy estimate, then report a confidence in the estimate, then finally compare to knowledge of result. In sports such as volleyball or basketball, significant differences exist from this paradigm. First, the outcome of the action is immediately known, thus it is not possible to separate the task from the knowledge of result. When a volleyball is spiked, the attacker immediately knows if it was a kill, in-play, or an error. When a basketball player shoots, it is either a basket or it is not, and the result is witnessed as soon as the ball leaves the shooter's hand. The same is true for a golfer who putts for the cup. The immediacy of outcome knowledge is a significant difference from previous calibration studies.

Second, the adversarial component in sport competition is different than typical calibration studies. In questions of knowledge, there was no opponent attempting to make the task more difficult for the subject. Card games, such as bridge, offer the variable of opponents to some degree. Once again, the type of sport may compare better than others. Even the type of event within the same sport may alter the applicability of the traditional paradigm. A track sprinter has opponents, but they are required to stay in their lanes and not interfere with other runners. Other track events do not have this restriction, and now the actions of opponents have a direct bearing on a runner's time. A runner may be boxed in among a group and unable to run at a familiar pace. In volleyball, several battles occur between the two teams. The server attempts to make the passer's job difficult, just as a blocker or digger attempts to prevent the attacker from scoring. The methods and tactics used to be successful are compounded by the methods and tactics used to prevent opponent's success. Adding another dimension to the already complex relationships, there are the intra-team dynamics that exist between teammates. While certain players can stand out from their teammates, these performances are also a result of the actions of teammates. The serve is the only task in volleyball that does not involve other players for execution. In basketball, the free throw similarly does not require interaction with any other player. However, unlike the free throw, the volleyball serve can be successful to a number of degrees. A serve can be easy or difficult for the opponent to pass, depending on such things as the speed, trajectory, final location of the serve, as well as the skill of the opponents. Therefore, being "correct" is not as clear as it was in previous calibration



studies. The calibration literature does not address the complex and dynamic systems at work in team, opponent referenced competitions.

### **Patterns of Calibration (R3)**

Finally, patterns of calibration were examined to determine if the expected overconfidence and hard-easy patterns emerge as seen in the literature, and whether these patterns hold for various levels of expertise. Unfortunately, distinctions of the overconfidence pattern were not clear. The exploratory nature of the graphical analyses suggested that the overconfidence pattern emerges if only considering strictly over-and underconfidence. However, when subjects who appear to be well-calibrated (closer to the ideal calibration line) are considered, the prevalence of overconfidence disappears. In this case, the results suggest that of the players who do not appear well-calibrated, more were underconfident. This result may again be indicative of the level of players within the sample. As developing players, the majority of which were low-performing in terms of performance cut points, lower confidence in estimates may be an acknowledgement of the lack of established and stable ability. Relating to previous calibration studies, it may be that the subjects acknowledge their lack of domain knowledge.

In determining whether the overconfidence pattern held with respect to expertise, results from the traditional comparison of confidence ratings to accuracy did not offer any conclusive support. However, when examining the sport-typical calibration bias, there was a clear trend for more expert players to be underconfident and less expert players to be overconfident. Put another way, the higher performing players had an easier time with the task while the poorer performing players had a more difficult time with the

task. This finding leads directly to examination of the hard-easy effect.

The tendency for underconfidence for easier tasks and overconfidence for harder tasks is the hard-easy effect. Because the high performers (easier tasks) were more underconfident, while the low performers (harder tasks) were more overconfident, the hard-easy pattern was found among the players. The significance of the pattern is unclear because, as Juslin and colleagues (2000) argue, the pattern may be an artifact of experimental method, rather than informational processing. The present study seems to fall into the category of the former. For example, if every player in the sample estimated the same performance level, those who performed better than this estimate would be underconfident, while those who performed below this estimate would be overconfident. The hard-easy effect would be evident, but only a methodological artifact.

**Perceived opponent difficulty.** The qualitative responses to the bases of perceived opponent difficulty offered insight into the reasons calibration is difficult to study in sport. In studies of cognitive ability, subjects are able to perform the task (i.e., answer the question) before making confidence judgments. Calibration studies in sport require estimation judgments before performing the task, because knowledge of result in executing a skill is immediate. This may not be the case in all sports, however. Track athletes, for example, are acutely aware of the time it takes them to cover specific distances. Swimmers are another example. Many of these athletes can be recruited based on verified times for certain events. Performance expectations would be accurate, and the development of the athlete would focus on reducing the time to cover a distance. In a sport such as volleyball, where multiple team and opponent variables coexist at every

skill execution, it is much more difficult to base performance outcome solely on one's own ability. The results of the qualitative measure seem to support the “noise plus bias” model by Burson et al., (2006) of general miscalibration in both skilled and unskilled performers. However, the finding that mastery-approach is correlated with calibration accuracy in the present study may indicate that calibration in volleyball can be improved if players make estimates based on understanding the limits of their own skill, rather than on external comparisons. However, the nature of the sport still requires an understanding of the inter- and intra-team influences on a player's capabilities.

In general, while the findings in the present study offer some support for previous literature regarding achievement goals and calibration, the lack of findings is arguably more significant. Achievement goal theory originated from educational classroom settings. Though some similarities exist, this domain is overall different from the sport domain. First, the students' overarching mission is not to “win” or “outperform” the other students. For athletes, this is an explicit goal to obtain. Second, the nature of classroom structure is different than the complex and dynamic structures inherent in team and opponent-referenced sports. Perhaps with the pressures of standardized testing, teachers may be able to learn from coaches, as the competitive and performance goals are beginning to emerge in the classroom. By making more explicit the fact that the “team” is competing against other “teams”, and the “coach” is evaluated by the team's performance may encourage motivation for studying.

Had achievement goal theory originated from the sport domain rather than the classroom, the measures may have evolved quite differently. Observation of behaviors

may have become the foundation for measuring goal orientation, rather than the remedial-looking, self-report questionnaires. In a similar vein, calibration in complex and dynamic systems such as team sports may have evolved calibration research in a much different way. The nature of the immediacy of outcome in many sports would have prohibited the simplistic comparison between beliefs of correct answers and actual correctness. Calibration would require not only knowledge of one's own abilities, but also of the abilities of those who seek to impede performance (opponents), the ability of those who seek to facilitate performance (teammates), and the interrelationships among them. For example, competing against a professional volleyball team would influence calibration differently than competing against a group of beginners. Likewise, competing against a professional team with professional teammates would influence calibration differently than competing against the same professional team with beginner teammates.

Both achievement goal theory and calibration research has evolved from more simple domains than the sport of volleyball. It is conceivable that the foundations of the early research no longer apply. Perhaps research in achievement goals and calibration would be better informed if the starting domain was the complex world of sport.

## **Implications**

Having a team whose members are highly motivated at all times is the ideal situation for any coach. The reality is that motivation levels vary not only among athletes, but also within the athletes themselves. Understanding how to maintain or encourage each player's motivation is a constant concern. At the least, coaches should be aware of what tasks and behaviors are demotivating for their players. Goal orientation theory offers a framework with which to identify types of motivation within each player. In the present study, two profiles of players emerged; high-all and high-approach. Both profiles of players would theoretically be motivated by practices aimed at improving or learning new skills as well as by competitive situations that offer the chance for them to be the best. The difference between the two would be seen in tasks that are aimed at maintaining already established skills and in competitive situations that do not identify the best performers. The high-approach profile would not be motivated to avoid becoming worse, nor to avoid being the worst. In practical terms, practices to reinforce fundamentals and competitions that are more group focused are not as motivating for the high-approach profile athlete.

The relationships between mastery-approach and calibration accuracy and between performance-approach and confidence in estimates are not surprising. Athletes who are oriented toward improving would have a clearer sense of their current ability,

and would thus display more accurate calibration. Athletes who are oriented toward proving their ability are likely to have had previous success that has reinforced their beliefs about themselves, thereby augmenting their overall confidence level. However, given the complexity and the variability inherent in the sport, calibration is not likely to be based solely on players' self-referenced competence. The most important question that a coach may have would be whether any of the orientations are indicative of better performers.

The lack of significant correlations with performance is disappointing from a coach's point of view. If a certain orientation were more indicative of higher performance, then coaches could recruit players with that orientation and potentially gain a performance advantage. It is interesting to note, however, the slight, negative correlations between performance and the avoid orientations. This result makes sense in that players (especially developing players) who are more motivated to avoid doing worse or being the worst are typically not ones who strive to improve and be the best. In a practical sense, coaches can modify practice and performance targets that would maximize each player's motivation to engage and persevere. For example, for those who are more approach-oriented, these targets would be to score as many points as possible in a given activity. The performance-approach players could compare their scores to the group, while the mastery-approach players could compare their scores to previous attempts. On the other hand, for those players who are avoid-oriented, targets would take the form of minimum thresholds. An example would be only those who perform above a certain level are able to move on to the next activity.

The initial expectation was that performance orientations, rather than mastery orientations, would be more prevalent due to the normative comparisons from competition. Although this expectation was not supported, the discrepancy can be explained from the perspective that collegiate athletes are closer to the developmental side of the ability continuum than to the established ability side. Therefore, the collegiate players are more oriented to developing than to performing. As a coach, this would be a significant realization. Coaches may be putting too much pressure on goals that are non-motivating for particular players, and may consequently damage the players' beliefs and feelings toward the game. For this reason, coaches should understand their role not only as performance directors, but also as educators. Player development requires a keen, working knowledge of the teaching-learning dynamic. This process is often overlooked, or altogether dismissed, by coaches who demand improved performance but are unable to help the players achieve it. The trend in many sports to specialize positions at an early age creates voids in skill sets, and limits the ability for these players to develop game fluency. As coaches, the need to teach in order to fill these voids is inevitable. Coaches must adapt with the changes in generations in order to develop players and teams with the desired ability to perform and successfully compete.

This trend is not universal, however. Different cultures choose to develop players more holistically and are prepared to sacrifice the winning now for the winning later. Some countries make it a rule for their domestic competition that the specific defensive player, called the "libero", cannot be used before a certain age. This encourages more players to develop all skills, rather than allowing players with undeveloped ball control

skills to be replaced by the libero in training as well as in matches. Of course, training players to develop all-around skills requires teaching and an understanding of encouraging motivation for improvement.



### **Limitations and Directions for Future Research**

One of the major limitations of the present study was the small sample size. As such, the sample and the corresponding results may not represent the population. Future studies should identify opportunities for similar studies with larger sample sizes. In conjunction with the sample size, the singular use of males represents a limitation. It would be interesting to see if gender differences exist using the same study. The use of questionnaires is also a limitation, as there may be some inclination to answer questions more in line with how subjects would like to be perceived, rather than how they actually are. Even with the adaptation of the AGQ-R for this study, the instrument may not be a good measure of achievement goals. Many of the items are worded similarly but with contradictory wording, which obscures and confuses the essence of the item. The vignettes offered an alternative measure that was ultimately more useful for cluster analysis. Future studies should continue to explore more effective methods of measuring goal orientations. For example, movement is the universal language in sport. Perhaps a better measure exists through direct observation of behaviors under scenarios that are set for specific goal orientations.

The study was also limited in the ability to distinguish between confidence in estimates used in the present study and sport-confidence, as discussed by Machida, Ward, and Vealey (2012). In their study, the authors extend Bandura's (1978) self-efficacy

model to sport, and define sport-confidence as “the beliefs in individual’s capability to be successful in sport” (Machida et al., 2012, p. 172). In short, the authors looked at antecedents to nine sources of sport-confidence among 206 collegiate athletes. These sources were bifurcated into controllable and uncontrollable sources. It was found that task-goal orientation (mastery-approach) and task-involving motivational climate were positive predictors of both controllable and uncontrollable sources, which fuel an enduring trait of sport-confidence. It is possible that participants in the present study are more accustomed to their sense of sport-confidence, rather than confidence in their estimates, and the confidence ratings reflect this. Therefore, calibration confidence results were a reflection of sport-confidence, rather than certainty of current ability.

That the initial expectation that performance orientations would dominate was not supported may be an indication of where the sample of players falls in the continuum of ability. Many of the best volleyball players in the world are significantly older and more experienced than collegiate players found in the United States. As such, collegiate players may be closer to the developmental levels than to performance levels. Cross sectional studies with established volleyball professionals would make for a more comprehensive analysis of goal-orientation profiles within the sport.

Along these lines, the next question of interest for future studies would be whether goal orientations change as players move along the ability continuum. While the cross sectional study suggested above would help to answer this question, it would be more telling to follow young, novice athletes through their career in a given sport and examine any changes in goal orientations that may occur. Additionally, understanding

whether overconfidence is a predictor for performance, or vice versa, would better inform coaches as to which area should be better supported. A longitudinal study may shed light on this by examining which begins to emerge first, performance or confidence, and may also inform on the pathways that lead to particular orientations. The value of a longitudinal study also lie in distinguishing between the context of training and competition. Goal orientations and perceived motivational climate have been shown to differ depending on whether the subjects answered questionnaires within the context of training or competition (van de Pol, Kavussanu, & Ring, 2012; van de Pol & Kavussanu, 2011). Armed with a better understanding of the contextual differences, sport practitioners may develop better methods of training and preparation for competition, which could enhance player development in both areas.

The use of existing calibration research to study calibration in sport is a significant limitation. Future calibration studies in sport should set aside calibration research as it currently stands and redefine what it means and how it is measured, based on the dynamic and complex systems at work in sports. One of the preliminary areas of focus should be identifying specific attributions of performance estimates. Various sports should be examined, ranging from the more individualized and isolated performances such as track and swimming, to the team and opponent-based competitions such as basketball and volleyball. Within specific sports, it would be beneficial to separate tasks into those that are considered closed (i.e., fixed variables) and open (i.e., dynamic variables). One of the fundamental differences between the two types of skills is how errors are defined. For example, errors in volleyball serving or basketball free throws can

be categorized as unforced, while errors in attacking or other shooting situations can be categorized as forced, due to the influence of the opponent challenging each action. Even within the unforced errors, there is influence of situational pressure that may influence errors in a less obvious, but equally powerful manner. Understanding these attributions may provide players, coaches, and sport practitioners with a compass to guide athlete development and competitive preparation.

Another area of calibration in sport to be studied would be the motivational effect of calibration accuracy. For example, if a player is well-calibrated at a certain performance level, what effect does this have on motivation to improve compared to the player who is poorly calibrated for the same level? Is there a difference between those who are undercalibrated and those who are overcalibrated in terms of motivation to improve? Armed with this type of information, coaches could prioritize activities that help players to calibrate themselves in order to increase motivation for subsequent training. In many ways coaches help players' calibration by giving them meaningful feedback in terms of process and outcome. Video review is also a powerful calibrator if used to help players reconcile what they believe they are doing with what they are actually doing. Creating a climate of critique without criticism can also facilitate calibration, especially if players are able to help each other with meaningful feedback. Coaches who choose their language carefully and actively seek to be a resource for improvement, rather than a judge of good and bad can create this type of climate.

## **Conclusion**

The results of the present exploratory study examined achievement goal orientation patterns and relationships with calibration accuracy and confidence, experience, and expertise in male, collegiate volleyball players. Profiles of goal orientations were identified that support the 2 x 2 Achievement Goal Framework (Elliot & McGregor, 2001a) and multiple goals theory (Pintrich, 2000). Relationships were found between mastery-approach and calibration accuracy, and between performance-approach and calibration confidence. The results showed mixed support for the general tendency for overconfidence in performance estimates. The hard-easy effect was found through graphical analysis of calibration bias, though the effect appears to be more of an artifact of methodology than of cognitive processing. Finally, qualitative analysis showed that calibration estimates were based on a variety of factors, many of which were external and uncontrollable.

## Appendices

## Appendix A

### Player Questionnaire Match 1

Team ID \_\_\_\_\_ Player ID \_\_\_\_\_

Please respond to each question on a scale of 1 (strongly disagree) to 5 (strongly agree).

With respect to my position...

|     |   |   |   |   |   |   |
|-----|---|---|---|---|---|---|
| 1.  | My aim is to completely master the skills I need.                 | 1 | 2 | 3 | 4 | 5 |
| 2.  | I strive to do well compared to other players.                    | 1 | 2 | 3 | 4 | 5 |
| 3.  | My goal is to improve as much as possible.                        | 1 | 2 | 3 | 4 | 5 |
| 4.  | My aim is to perform well relative to other players.              | 1 | 2 | 3 | 4 | 5 |
| 5.  | My aim is to avoid getting worse as a player.                     | 1 | 2 | 3 | 4 | 5 |
| 6.  | My goal is to avoid performing poorly compared to other players.  | 1 | 2 | 3 | 4 | 5 |
| 7.  | I strive to understand my position as thoroughly as possible.     | 1 | 2 | 3 | 4 | 5 |
| 8.  | My goal is to perform better than the other players.              | 1 | 2 | 3 | 4 | 5 |
| 9.  | My goal is to avoid playing worse than I played in my last match. | 1 | 2 | 3 | 4 | 5 |
| 10. | I strive to avoid performing worse than other players.            | 1 | 2 | 3 | 4 | 5 |

|     |   |   |   |   |   |   |
|-----|---|---|---|---|---|---|
| 11. | I strive to avoid an incomplete understanding of my position. | 1 | 2 | 3 | 4 | 5 |
| 12. | My aim is to avoid doing worse than other players.            | 1 | 2 | 3 | 4 | 5 |



Please read the following 4 descriptions and indicate how much each is like you based on a scale of 1 (not at all like me) to 5 (very much like me)

1. I like knowing how I played compared to the other players. I work hard because I do not want to be the worst on the team. I practice harder when my coach posts everyone's performance. My performance has more to do with my ability than my effort. If I play in a match, I look at match statistics to see how I did compared to everyone else. If I don't perform well, it's usually because the other players were better than me.

circle your answer

|                    |   |   |   |               |
|--------------------|---|---|---|---------------|
| 1                  | 2 | 3 | 4 | 5             |
| not at all like me |   |   |   | a lot like me |

2. I like knowing that I did not play worse than my previous performance. I work hard to make sure my skills do not deteriorate. It makes no difference to me if my coach posts everyone's performance in practice. My performance has more to do with my effort than my ability. If I play in a match, I look at match statistics to see how I did compared to my last performance. If I don't perform well, it's usually because I did not work hard enough.

circle your answer

|                    |   |   |   |               |
|--------------------|---|---|---|---------------|
| 1                  | 2 | 3 | 4 | 5             |
| not at all like me |   |   |   | a lot like me |

3. I like knowing how I played compared to the other players. I work hard because I want to be the best on my team. I practice harder when my coach posts everyone's performance. My performance has more to do with my ability than my effort. If I play in a match, I look at match statistics to see how I did compared to everyone else. If I don't perform well, it's usually because I was unlucky.

circle your answer

|                    |   |   |   |               |
|--------------------|---|---|---|---------------|
| 1                  | 2 | 3 | 4 | 5             |
| not at all like me |   |   |   | a lot like me |

4. I like knowing if I have improved compared to my previous performance. I work hard to be better than I was before. It makes no difference to me if my coach posts everyone's performance in practice. My performance has more to do with my effort than my ability. If I play in a match, I look at match statistics to see how I did compared to my last performance. If I don't perform well, it's usually because I did not prepare well.

circle your answer

1                      2                      3                      4                      5

not at all like me

a lot like me

Please answer the following questions.

1. What is your current age? \_\_\_\_\_ Years

2. What is your ethnicity?

\_\_\_\_\_  
Black/African  
American

\_\_\_\_\_  
White/Caucasian

\_\_\_\_\_  
Asian

\_\_\_\_\_  
Other

\_\_\_\_\_  
Mixed (please  
write in primary  
two)

3. What is your academic class designation? (mark with X)

\_\_\_\_\_  
Freshman

\_\_\_\_\_  
Sophomore

\_\_\_\_\_  
Junior

\_\_\_\_\_  
Senior

\_\_\_\_\_  
Fifth Year

4. What NCAA Division designation is your program? (circle one)  
III

I

II

5. How many years have you played volleyball at the following levels? (exclude current year)

\_\_\_\_\_  
High School

\_\_\_\_\_  
Club

\_\_\_\_\_  
USA High  
Performance

\_\_\_\_\_  
College

6. Based on previous performance, were you given any specific performance targets for your upcoming match? (circle: yes / no) If “yes”, what are they?

7. How would you rate the ability of your upcoming opponent? (mark with X)

\_\_\_\_\_  
Low

\_\_\_\_\_  
Somewhat Low

\_\_\_\_\_  
Average

\_\_\_\_\_  
Somewhat High

\_\_\_\_\_  
High

8. What is this rating based on?

Please answer the following questions as accurately as possible. Estimates should be realistic and not what you would like them to be.

**Attacking**

9. What do you estimate will be your percentages among the following categories for YOUR ATTACKS?

|                               |                 |                 |                 |                 |                               |
|-------------------------------|-----------------|-----------------|-----------------|-----------------|-------------------------------|
| (total should add up to 100%) | <u>      </u> % | <u>      </u> % | <u>      </u> % | <u>      </u> % | <u>      </u> %               |
|                               | Kill for point  | In play         | Blocked         | Error           | I do not attack (mark with X) |

10. How confident are you in your estimates?  
      %

11. What do you estimate will be your percentages among the following categories for your TEAM ATTACKS?

|                               |                 |                 |                 |                 |
|-------------------------------|-----------------|-----------------|-----------------|-----------------|
| (total should add up to 100%) | <u>      </u> % | <u>      </u> % | <u>      </u> % | <u>      </u> % |
|                               | Kill for point  | In play         | Blocked         | Error           |

12. How confident are you in your estimates?  
      %

## Questionnaire Match 2

Team ID \_\_\_\_\_ Player ID \_\_\_\_\_

1. Based on previous performance, were you given any specific performance targets for your upcoming match? (circle: yes / no) If “yes”, what are they?

2. How would you rate the ability of your upcoming opponent? (mark with ✕)

\_\_\_\_\_ Low      \_\_\_\_\_ Somewhat Low      \_\_\_\_\_ Average      \_\_\_\_\_ Somewhat High      \_\_\_\_\_ High

3. What is this rating based on?

Please answer the following questions as accurately as possible. Estimates should be realistic and not what you would like them to be.

### Attacking

4. What do you estimate will be your percentages among the following categories for YOUR ATTACKS?

(total should add up to 100%)      \_\_\_\_\_% Kill for point      \_\_\_\_\_% In play      \_\_\_\_\_% Blocked      \_\_\_\_\_% Error      \_\_\_\_\_% I do not attack (mark with ✕)

5. How confident are you in your estimates?

\_\_\_\_\_%

6. What do you estimate will be your percentages among the following categories for your TEAM ATTACKS?

(total should add up to 100%)      \_\_\_\_\_% Kill for point      \_\_\_\_\_% In play      \_\_\_\_\_% Blocked      \_\_\_\_\_% Error

7. How confident are you in your estimates?

\_\_\_\_\_%

### Questionnaire Match 3

Team ID \_\_\_\_\_ Player ID \_\_\_\_\_

1. Based on previous performance, were you given any specific performance targets for your upcoming match? (circle: yes / no) If “yes”, what are they?

2. How would you rate the ability of your upcoming opponent? (mark with ✕)

\_\_\_\_\_ Low      \_\_\_\_\_ Somewhat Low      \_\_\_\_\_ Average      \_\_\_\_\_ Somewhat High      \_\_\_\_\_ High

3. What is this rating based on?

Please answer the following questions as accurately as possible. Estimates should be realistic and not what you would like them to be.

#### Attacking

4. What do you estimate will be your percentages among the following categories for YOUR ATTACKS?

(total should add up to 100%)      \_\_\_\_\_% Kill for point      \_\_\_\_\_% In play      \_\_\_\_\_% Blocked      \_\_\_\_\_% Error      \_\_\_\_\_% I do not attack (mark with ✕)

5. How confident are you in your estimates?

\_\_\_\_\_%

6. What do you estimate will be your percentages among the following categories for your TEAM ATTACKS?

(total should add up to 100%)      \_\_\_\_\_% Kill for point      \_\_\_\_\_% In play      \_\_\_\_\_% Blocked      \_\_\_\_\_% Error

7. How confident are you in your estimates?

\_\_\_\_\_%

## Appendix B

### Match Data Sheet

TEAM ID \_\_\_\_\_

| Match #1  |             |               |               |              |
|-----------|-------------|---------------|---------------|--------------|
| Player ID | Total Kills | Total In-Play | Total Blocked | Total Errors |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
| Match #2  |             |               |               |              |
| Player ID | Total Kills | Total In-Play | Total Blocked | Total Errors |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
| Match #3  |             |               |               |              |
| Player ID | Total Kills | Total In-Play | Total Blocked | Total Errors |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |
|           |             |               |               |              |

## Appendix C

### CALIBRATION AND ACHIEVEMENT GOALS IN COLLEGE VOLLEYBALL

#### **INFORMED CONSENT FORM**

##### **RESEARCH PROCEDURES**

This research is being conducted to examine achievement goal orientation patterns and relationships with calibration accuracy and confidence, expertise, practice, feedback, and expected opponent (task) difficulty, in competitive, collegiate volleyball players at NCAA member universities.

Coaches. If you agree to participate, you will be asked to create a player-letter key, and assign letters to each of your players. The key will not be seen or collected by the researcher. You will also be asked to keep player attack data for three matches that you designate for the purpose of this study. The letters will be used to match player questionnaires with match data. If you choose not to participate, your players will not be asked to participate.

Players. If you agree to participate, you will be asked to fill out a questionnaire prior to an upcoming match on three separate occasions. The first questionnaire will take approximately 25 minutes, while the remaining two will take approximately 5 minutes each. Your names will not be used throughout the study. You will use a player identification letter assigned to you by your coach. Your questionnaires will be linked to your match data that will be provided to the researcher by your coach.

##### **RISKS**

There are no foreseeable risks for participating in this research.

##### **BENEFITS**

There are no benefits to you as a participant other than to further research in calibration and achievement goal theory.

##### **CONFIDENTIALITY**

The data in this study will be confidential. Identification letters will be issued to each participant and the following will be adhered: Identification letters will be issued to each participant by their coach prior to any data collection and the following will be adhered: (1) your name will not be included on the surveys and other collected data; (2) your identification letter will be placed on the survey and other collected data; (3) through the use of an identification letter, the researcher



will be able to link your survey to your match data; and (4) the researcher will not have access to the identification key.

### **PARTICIPATION**

Your participation is voluntary, and you may withdraw from the study at any time and for any reason. If you decide not to participate or if you withdraw from the study, there is no penalty or loss of benefits to which you are otherwise entitled. There are no costs to you or any other party

### **CONTACT**

This research is being conducted by Fred Chao, Department of Educational Psychology at George Mason University. He may be reached at 703-993-3227 for questions or to report a research-related problem. The faculty advisor is Dr. Angela Miller, who can be reached at 703-993-5590. You may contact the George Mason University Office of Research Integrity & Assurance at 703-993-4121 if you have questions or comments regarding your rights as a participant in the research.

This research has been reviewed according to George Mason University procedures governing your participation in this research.

### **CONSENT**

I have read this form and agree to participate in this study.

---

Signature

---

Name

---

Date of Signature

## References

- Alexander, P. A. (2013). Calibration: What is it and why it matters? An introduction to the special issue on calibrating calibration. *Learning and Instruction*, 24, 1–3. doi:10.1016/j.learninstruc.2012.10.003
- Ames, C. (1984). Achievement attributions and self-instructions under competitive and individualistic goal structures. *Journal of Educational Psychology*, 76(3), 478–487. doi:10.1037/0022-0663.76.3.478
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84(3), 261–271. doi:10.1037/0022-0663.84.3.261
- Amorose, A. J., & Anderson-Butcher, D. (2007). Autonomy-supportive coaching and self-determined motivation in high school and college athletes: A test of self-determination theory. *Advances in Self-Determination Theory Research in Sport and Exercise*, 8(5), 654–670. doi:10.1016/j.psychsport.2006.11.003
- Amorose, A. J., & Horn, T. S. (2000). Intrinsic motivation: Relationships with collegiate athletes' gender, scholarship status, and perceptions of their coaches' behavior. *Journal of Sport & Exercise Psychology*, 22(1), 63–84.
- Amorose, A. J., & Horn, T. S. (2001). Pre- to Post-Season Changes in the Intrinsic Motivation of First Year College Athletes: Relationships with Coaching Behavior

- and Scholarship Status. *Journal of Applied Sport Psychology*, 13(4), 355–373.  
doi:10.1080/104132001753226247
- Arkes, H. R., Christensen, C., Lai, C., & Blumer, C. (1987). Two methods of reducing overconfidence. *Organizational Behavior and Human Decision Processes*, 39(1), 133–144. doi:10.1016/0749-5978(87)90049-5
- Bandura, A. (1978). Self-efficacy: Toward a unifying theory of behavioral change. *Advances in Behaviour Research and Therapy*, 1(4), 139–161. doi:10.1016/0146-6402(78)90002-4
- Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology*, 90(1), 60–77.  
doi:10.1037/0022-3514.90.1.60
- Cervelló, E. M., & Santos-Rosa, F. J. (2001). Motivation in sport: An achievement goal perspective in young Spanish recreational athletes. *Perceptual and Motor Skills*, 92(2), 527–534.
- Christensen-Szalanski, J. J., & Bushyhead, J. B. (1981). Physicians' use of probabilistic information in a real clinical setting. *Journal of Experimental Psychology: Human Perception and Performance*, 7(4), 928–935. doi:10.1037/0096-1523.7.4.928
- Conroy, D. E., Elliot, A. J., & Hofer, S. M. (2003). A 2 x 2 achievement goals questionnaire for sport: evidence for factorial invariance, temporal stability, and external validity. *Journal of Sport & Exercise Psychology*, 25(4), 456–476.

- Cury, F. (2004). Évolution conceptuelle de la théorie des buts d'accomplissement dans le domaine du sport. [Theoretical advances in achievement goal theory in sport setting.]. *L'Année Psychologique*, 104(2), 295–330. doi:10.3406/psy.2004.29669
- Daniels, L. M., Haynes, T. L., Stupnisky, R. H., Perry, R. P., Newall, N. E., & Pekrun, R. (2008). Individual differences in achievement goals: A longitudinal study of cognitive, emotional, and achievement outcomes. *Contemporary Educational Psychology*, 33(4), 584–608. doi:10.1016/j.cedpsych.2007.08.002
- Duda, J. L., & Nicholls, J. G. (1992). Dimensions of achievement motivation in schoolwork and sport. *Journal of Educational Psychology*, 84(3), 290–299. doi:10.1037/0022-0663.84.3.290
- Dweck, C. S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *Journal of Personality and Social Psychology*, 31(4), 674–685. doi:10.1037/h0077149
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41(10), 1040–1048. doi:10.1037/0003-066X.41.10.1040
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2), 256–273. doi:10.1037/0033-295X.95.2.256
- Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34(3), 169–189. doi:10.1207/s15326985ep3403\_3

- Elliot, A. J., & McGregor, H. A. (2001a). A  $2 \times 2$  achievement goal framework. *Journal of Personality and Social Psychology*, 80(3), 501–519. doi:10.1037/0022-3514.80.3.501
- Elliot, A. J., & McGregor, H. A. (2001b). A  $2 \times 2$  achievement goal framework. *Journal of Personality and Social Psychology*, 80(3), 501.
- Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, 100(3), 613–628. doi:10.1037/0022-0663.100.3.613
- Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54(1), 5–12. doi:10.1037/0022-3514.54.1.5
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 3(4), 552–564. doi:10.1037/0096-1523.3.4.552
- Fogarty, G. J., & Else, D. (2005). Performance calibration in sport: Implications for self-confidence and metacognitive biases. *International Journal of Sport and Exercise Psychology*, 3(1), 41–57. doi:10.1080/1612197X.2005.9671757
- Fogarty, G. J., & Ross, A. (2007, September 25). Calibration in tennis: the role of feedback and expertise. *Proceedings of the 42nd Australian Psychological Society Annual Conference: Psychology Making an Impact*. Retrieved October 10, 2012, from [http://www.psychology.org.au/publications/conference\\_proceedings/](http://www.psychology.org.au/publications/conference_proceedings/)

- Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard–easy effect. *Psychological Review*, 107(2), 384–396. doi:10.1037/0033-295X.107.2.384
- Keren, G. (1987). Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes*, 39(1), 98–114. doi:10.1016/0749-5978(87)90047-1
- Keren, G. (1991). Calibration and probability judgements: Conceptual and methodological issues. *Acta Psychologica*, 77(3), 217–273. doi:10.1016/0001-6918(91)90036-Y
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 107–118. doi:10.1037/0278-7393.6.2.107
- Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General*, 131(2), 147–162. doi:10.1037/0096-3445.131.2.147
- Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance*, 20(2), 159–183. doi:10.1016/0030-5073(77)90001-0
- Lichtenstein, S., & Fischhoff, B. (1980). Training for calibration. *Organizational Behavior and Human Performance*, 26(2), 149–171. doi:10.1016/0030-5073(80)90052-5

- Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1981). *Calibration of Probabilities: The State of the Art to 1980*. DTIC Document.
- Machida, M., Ward, R. M., & Vealey, R. S. (2012). Predictors of sources of self-confidence in collegiate athletes. *International Journal of Sport and Exercise Psychology*, 10(3), 172–185.
- Nicholls, J. G. (1978). The Development of the Concepts of Effort and Ability, Perception of Academic Attainment, and the Understanding That Difficult Tasks Require More Ability. *Child Development*, 49(3), 800–814. doi:10.2307/1128250
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91(3), 328–346. doi:10.1037/0033-295X.91.3.328
- Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92(3), 544–555. doi:10.1037/0022-0663.92.3.544
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1), 54–67. doi:10.1006/ceps.1999.1020
- Seifriz, J. J., Duda, J. L., & Chi, L. (1992). The relationship of perceived motivational climate to intrinsic motivation and beliefs about success in basketball. *Journal of Sport & Exercise Psychology*, 14(4), 375–391.
- Smith, A. L., Balaguer, I., & Duda, J. L. (2006). Goal orientation profile differences on perceived motivational climate, perceived peer relationships, and motivation-

- related responses of youth athletes. *Journal of Sports Sciences*, 24(12), 1315–1327. doi:10.1080/02640410500520427
- Smith, R. E., Smoll, F. L., & Cumming, S. P. (2009). Motivational climate and changes in young athletes' achievement goal orientations. *Motivation and Emotion*, 33(2), 173–183. doi:10.1007/s11031-009-9126-4
- Van de Pol, P. K. C., & Kavussanu, M. (2011). Achievement goals and motivational responses in tennis: Does the context matter? *Psychology of Sport and Exercise*, 12(2), 176–183. doi:10.1016/j.psychsport.2010.09.005
- Van de Pol, P. K. C., Kavussanu, M., & Ring, C. (2012). Goal orientations, perceived motivational climate, and motivational outcomes in football: A comparison between training and competition contexts. *Psychology of Sport and Exercise*, 13(4), 491–499. doi:10.1016/j.psychsport.2011.12.002
- Wahlstrom, D. A. (2001). *Technology-mediated assessment in a secondary manufacturing technology program: A study of the relationship between participants' goal-orientation and the calibration of performance expectations to performance feedback* (Ph.D.). University of Michigan, United States -- Michigan. Retrieved from <http://search.proquest.com.mutex.gmu.edu/pqdtft/docview/275921231/abstract/1422C1415F87AD9C4B4/1?accountid=14541>
- Wang, C. K. J., Biddle, S. J. H., & Elliot, A. J. (2007). The 2×2 achievement goal framework in a physical education context. *Psychology of Sport and Exercise*, 8(2), 147–168. doi:10.1016/j.psychsport.2005.08.012





## Biography

Fred Chao graduated from Paint Branch High School, Burtonsville, Maryland, in 1990. He received his Bachelor of Science from George Mason University in 1994. He has been employed as the head coach for the George Mason University Men's Volleyball program since 1998.