

Fall 2006

Coach Effectiveness and Personality Assessments: An Exploratory Analysis of Thin Slice Interpersonal Perceptions

Jennifer Lynne Knight

Follow this and additional works at: <https://digitalcommons.georgiasouthern.edu/etd>

Recommended Citation

Knight, Jennifer Lynne, "Coach Effectiveness and Personality Assessments: An Exploratory Analysis of Thin Slice Interpersonal Perceptions" (2006). *Electronic Theses and Dissertations*. 75.

<https://digitalcommons.georgiasouthern.edu/etd/75>

This thesis (open access) is brought to you for free and open access by the Graduate Studies, Jack N. Averitt College of at Digital Commons@Georgia Southern. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons@Georgia Southern. For more information, please contact digitalcommons@georgiasouthern.edu.

COACH EFFECTIVENESS AND PERSONALITY ASSESSMENTS: AN EXPLORATORY ANALYSIS OF THIN SLICE INTERPERSONAL PERCEPTIONS

by

JENNIFER L. KNIGHT

(Under the Direction of Daniel R. Czech)

ABSTRACT

This research examined accuracy, consensus, and self-other agreement of personality assessments and coaching effectiveness based on thin-slice judgments of 30-second video clips of 9 recreation level coaches. Two hundred and six naïve raters viewed the clips and rated the targets on coaching effectiveness and personality attributes. Ratings of coaching effectiveness were correlated with expert ratings of effectiveness to measure accuracy. The ratings of attributes were correlated with expert ratings of the same attributes to measure consensus. Gender, race, and level of sport participation of naïve raters was subjected to independent samples t-tests and one-way analyses of variance (ANOVA) to determine if they moderated thin-slice judgments. Results indicated that naïve raters as a group were not accurate in assessment of coaching effectiveness, nor were there significant correlations on consensus or self-other agreement. There were significant differences between levels of sport participation groups on two of the fourteen attributes: competence and confidence.

INDEX WORDS: Thin-slicing, Coaching Effectiveness, Consensus, Accuracy

COACH EFFECTIVENESS AND PERSONALITY ASSESSMENTS: AN
EXPLORATORY ANALYSIS OF THIN SLICE INTERPERSONAL PERCEPTIONS

by

JENNIFER LYNNE KNIGHT

B.S., University of Akron, 2001

A thesis Submitted to the Graduate Faculty of Georgia Southern University
in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

STATESBORO, GEORGIA

2006

© 2006

Jennifer Lynne Knight

All Rights Reserved

ELECTRONIC THESIS APPROVAL PAGE

by

JENNIFER LYNNE KNIGHT

Major Professor:

Daniel R. Czech

Committee:

A. Barry Joyner
Jonathan N. Metzler
Alan D. Zwald
Larry D. Bryant

Electronic Version Approved:
August 2006

DEDICATION

I would like to dedicate this project to three people in my life who are probably happier than I am about this accomplishment. First to my parents who always believed I was smarter and more talented than I ever did! They have always supported me despite my very long circuitous path to this place in life.

My best friend Renee who strives every day to be a better person and reach the goals she has set. She has always been there to encourage me and keep me moving forward. Her phone calls from Croatia during my all-nighters were welcome distractions!

ACKNOWLEDGEMENTS

I would like to thank my committee for their guidance throughout this project: Dr. Czech for the idea; Coach Bryant for the openness and use of his students; Dr. Joyner for his consistency and kindness; Dr. Zwald for his expertise, and Jonathan Metzler for his open door and sharp research skills.

I would also like to thank many people in my life without whose support I would not be where I am today. First of all to the rest of my sport psych class: Sarah, Mike, Del, Karen, and Marty- we have certainly shared many ups and downs over our two years! I have learned from each one of you and I appreciate all of the constructive criticism I have received in the many classes we have shared together. I wish you all the best! Of course, thanks to my office buddy Vinny who was always around to listen!

To all my friends at GSU who made the 'Boro fun and worthwhile! Of course, thanks to my family.

Lastly, I would like to thank everyone who supports me and enjoys me despite my unconventional ways!

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	6
LIST OF TABLES.....	8
CHAPTER	
1. INTRODUCTION.....	9
Statement of Purpose.....	14
2. METHODS.....	16
Participants.....	16
Instrumentation.....	16
Procedures.....	17
Data Analysis.....	18
3. RESULTS.....	20
4. DISCUSSION.....	26
REFERENCES.....	36
APPENDICES.....	44
A RESEARCH QUESTIONS, LIMITATIONS, DELIMITATIONS, ASSUMPTIONS, DEFINITIONS OF TERMS.....	45
B EXTENDED REVIEW OF LITERATURE, ADDITIONAL REVIEW OF LITERATURE REFERENCES.....	50
C INSTRUMENTATION.....	69

LIST OF TABLES

	Page
Table 1: Descriptive Statistics.....	21
Table 2: Descriptive Statistics for Target Accuracy and Consensus Differentiated by Gender.....	23
Table 3: Descriptive Statistics for Target Accuracy and Consensus Differentiated by Race.....	24
Table 4: Analysis of Variance for Attributes between Levels of Sport Participation	25

CHAPTER 1

INTRODUCTION

Gordon Allport (1937) suggested that people are able to form accurate perceptions of others from mere glimpses of their behavior. Making judgments from so called “thin slices” of behavior has become very popular in contemporary social psychological research (Ambady, Hallahan, & Rosenthal, 1995; Ambady & Gray, 2002; Ambady & Rosenthal, 1992, 1993). The concept of interpersonal perception accuracy based solely on thin slices has been brought to mainstream attention by the popular book *Blink* by Malcolm Gladwell that suggests that most people can thin-slice with surprising success, so that “decisions made very quickly can be as good as decisions made consciously and deliberately (Gladwell, 2005, p. 14)”. Gladwell provides examples from academic research to support his overall premise, including that of Ambady and Rosenthal (1993). Thin-slices are brief excerpts of expressive behavior less than five minutes sampled from the behavioral stream (Ambady & Gray, 2002).

Ambady and Rosenthal (1992) suggested that expressive behaviors (movement, speech, gesture, facial expressions, posture) contribute to impressions made about the target. Early researchers were interested in the link between expressive behaviors as the indicators of personality (Allport, 1937; Allport & Vernon, 1933). The cues that are projected by expressive behavior have been shown to be interpreted accurately in as little as a 2-second nonverbal clip of a target (Ambady & Rosenthal, 1993).

Ambady and Rosenthal (1992) suggested that the accuracy of thin-slice judgments have practical applications in fields that are interpersonally oriented. When thin slice ratings predict criterion variables, they can be used, for example, to target biased teachers

or gauge expectancies of newscasters. They also suggest that thin slice judgments can be used in the selection, training, and evaluation of people in fields where interpersonal skills are important. Accuracy of thin-slice judgments of coaches could be very useful in selection, training, and evaluation of coaches.

Accuracy in personality and social psychology research can be defined in three ways: the degree of correspondence between a judgment and a criterion; interpersonal consensus; and a construct possessing pragmatic utility (Kruglanski, 1989). These definitions fall into two approaches within the field. The pragmatic approach defines a judgment as accurate if it predicts behavior. This approach looks at personality judgments as necessary tools for social living and evaluates their accuracy in terms of their practical value (Funder, 1995). The constructivist approach focuses on consensus between raters. This approach looks at all judgments as perceptions and evaluates their accuracy in terms of agreement between judges (Funder, 1995). Kenny (1994) further explained that target accuracy is broken into three categories: perceiver, generalized, and dyadic. Generalized target accuracy is the correlation between how a person is generally seen by others and how that person generally behaves. Target accuracy can be defined in thin-slice research as the correspondence between participants' judgments of a target individual and well-defined external criterion (Ambady & Hall, 2002; Ambady & Rosenthal, 1992, 1993).

Thin-slice judgments have been shown to produce similar judgments to ecologically valid criterion. Ecologically valid criteria are characterized by pragmatic utility in that they are used in everyday decisions about people as an external outcome of observed behavior (Ambady & Rosenthal, 1993). Support for congruence in this relationship has been shown by significant positive correlations between naïve judgments

and outcomes, such as predicting judgments of candidates in job interviews and effectiveness of teachers (Ambady, Hallahan, & Rosenthal, 1995).

The target accuracy and consensus of naïve raters given thin-slices of information appears to be moderated by characteristics of the raters, traits assessed, and characteristics of the targets. Studies show that individual differences of raters can affect judgments based on thin-slices of information including gender and ethnicity (Ambady, & Gray, 2002; Ambady, Hallahan, & Rosenthal, 1995, Fortunato, & Mincy, 2003; Vrij, Dragt, & Koppelaar, 1992). Previous research is equivocal regarding the accuracy of judgments based on gender. Some research suggests that females are more accurate judges of non-verbal behavior (Hall, 1984), while other research found no difference in judgments of non-verbal behavior based on gender (Ambady & Rosenthal, 1992). Researchers have found that raters judge targets of a different ethnicity more negatively than targets of the same ethnicity (Vrij, Dragt, & Koppelaar, 1992).

Another bias can involve the dimensions being rated. One study found accuracy at zero acquaintance for judgments of extraversion, but not conscientiousness (Kenny, Horner, Kashy, & Chu, 1992). Another study found similar correlations for extraversion as well as a relationship between zero acquaintance ratings of conscientiousness, but not for agreeableness, emotional stability, and culture (Borkenau & Liebler, 1992). John and Robins (1993) suggest that differences in ratings on traits depend on evaluativeness and observability. Traits that are less evaluative (neutral) and more observable reach greater consensus and accuracy (John and Robins, 1993). They define observability by the degree to which behaviors are relevant to the trait can be easily observed. They define evaluativeness by the degree to which a trait is relatively neutral.

Limitations are also present on the persons being judged. Persons who possess extraversion and good mental health are simpler to judge at first glance than targets who possess introversion or poor mental health, because “exterior behavior mimics their internal view of themselves. What you see is what you get” (Flora, 2004, p. 66). Social context can also play a role depending on personality types. Expressive behaviors were limited by individuals with a high self-monitoring in social situations, therefore making judgments on their mood more difficult (Friedman Miller-Herringer, 1991).&

Ambady and Rosenthal (1993) researched intuitive judgments on teacher effectiveness. It was determined that thin-slice evaluations by naive raters of 30 seconds, 5 seconds, and 2 seconds were congruent with evaluations by students and principals who observed the teacher for a semester. It is suggested that the accuracy of the thin-slice judgments can be attributed to raters’ years experience in classroom situations; therefore, within the coaching context, amount of sport experience may also be an individual difference that moderates interpersonal perception accuracy. Ambady and Rosenthal measured judgments on fourteen personality attributes: accepting, active, dominant, empathic, enthusiastic, honest, likable, optimistic, professional, supportive, and warm. Teaching is an interpersonal field, as well as coaching. Due to similarities in the fields the same attributes were chosen in this study.

The teaching and coaching environment may have parallels and crossover applications. Often cited in coaching and teaching lore is John Wooden, who was one of the most successful collegiate basketball coaches. Wooden pointed out that coaches are teachers first and profiled ten criteria needed for a successful teacher; among them, knowledge and warm personality and genuine consideration of others (Wooden, 1980).

Research in the teaching profession highlights attributes of successful teaching. The list includes a teacher's enthusiasm and positive attitude, approachability, an environment that is positive, cooperative, and clear-cut, specific objectives, as well as appropriate feedback (Csikszentmihalyi & McCormack, 1986; McGinnis, 1985; Slavin, 1980). Wooden's coaching philosophy includes all of the aforementioned in his pyramid of success. Bloom (1985) explains that coaching, like teaching, can perhaps best be viewed as an interpersonal relations field, which rests primarily on effective communication and interaction among various participants.

Coaching research has identified behaviors that elicit positive perceptions from athletes towards coaches (Smith & Smoll, 1997, Williams et al, 2003). Behaviors include positive reinforcement, technical instruction, encouragement, and structuring fun practices. It is theorized that a coach's behavior plays a significant role in the psychological development of young athletes (Smith & Smoll, 1990). Youth sport research highlights the positive relationship between specific coaching behaviors and self-esteem, satisfaction, and enjoyment in children (Smith & Smoll, 1990; Smoll, Smith, Barnett, & Everett, 1993). This has led to a recent theoretical model (Conroy & Coatsworth, 2006) that emphasizes how coaching behaviors impact youth psychosocial outcomes which emphasizes the role of athletes' perceptions.

A recent study explored the characteristics of expert university level coaches and found several personal attributes that these coaches possessed: commitment to learning; learning from past mistakes; knowledgeable; open-minded; balanced; composed; caring; and genuinely interested in their athletes (Vallee & Bloom, 2005).

Previous research targets the importance of increasing self-awareness of coaches' referencing personal behavior while coaching (Smith & Smoll, 1997; Smith, Smoll, & Curtis, 1978; Wandzilak, Ansorge, & Potter, 1988). In a study that coded coaches' behaviors, the athletes were significantly more successful than the coaches' in the recall of those behaviors (Smith & Smoll, 1997). This same research determined that youth athletes' interpretation of coaches' behaviors are of even greater impact than the actual behaviors in psychosocial outcomes. At the recreation level, game outcomes bear little significance in psychosocial outcomes (reaction to coach, enjoyment, and self-esteem) for the athletes. The measurement of psychosocial outcomes showed a significant relationship between coaches' behavior and aforementioned outcomes. Earlier research (Bloom, 1985) indicated that the coach is central to the development of expertise in a sport.

Nonverbal behavior can be very significant in an environment where high levels of stress and decision-making are concerned. Perceptions can cause shifts in confidence. Research supports that the self-efficacy of athletes who judged opponents non-verbal behavior was directly related to those perceptions (Greenlees, Bradley, Holder, & Thelwell, 2005). As outcome expectations may be influenced by perceptions of sporting opponents, and have been shown to influence performance levels (Eyal, Bar-Eli, Tenenbaum, & Pie, 1995; Feltz & Riessinger, 1990; Weinberg, Gould, & Jackson, 1979).

Statement of Purpose

Therefore, the purpose of this study is to examine the relationship between naïve ratings of thin-slices of coaching and ecologically valid criterion measures, which are end of the season evaluations by supervisors, as well as self measures of coaching attributes

and effectiveness. This research will also include the demographic background of the naïve raters and explore the differences among evaluations based on gender, race, and level of sport participation. The following nine research questions are explored: What is the naïve raters' accuracy in their assessment of coaching effectiveness; what is the consensus between naïve raters and experts on each attribute; what is the self-other agreement between naïve raters and coaches on each attribute; is there a significant difference in accuracy between male and female raters; is there a significant difference in consensus between male and female raters; is there a significant difference in accuracy between races of raters; is there a significant difference in consensus between races of raters; is there a significant difference in accuracy between raters' level of sport participation; is there a significant difference in consensus between raters' level of sport participation?

CHAPTER 2

METHODS

Participants

There were two samples of participants in this study. Sample A consisted of 206 naïve raters recruited from undergraduate healthful living classes. Raters ranged from 18 to 55 years old ($M = 19.6$; $SD = 4.4$) and included 115 men and 91 women. Raters included African-Americans ($n = 47$), Caucasians ($n = 147$), Hispanics ($n = 6$), and other races ($n = 4$). The naïve raters indicated the highest level of sport in which they participated: none ($n = 26$); recreation ($n = 46$); junior varsity ($n = 16$); varsity/elite ($n = 91$); and college ($n = 20$). Sample B consisted of nine coaching students (eight men, one woman) from an undergraduate level coaching course at a southeastern university. There were eight Caucasian coaches and one African-American coach. The average age of the coaches was 20.2 years old ($SD = 1.4$)

Instrumentation

Coach attributions. Naïve raters, coaches, and supervisors rated each coach using an attributional survey (Ambady & Rosenthal, 1993) which included the following subscales: accepting, active, dominant, empathic, enthusiastic, honest, likable, optimistic, professional, supportive, and warm. Each coach was rated three times for each attribute on a 9-point Likert scale ranging from *not at all* (1) to *very* (9). The reliability in previous research of the mean of the judges' ratings of the sum of the mean ratings of the 14 nonverbal variables was .80 (Ambady & Rosenthal, 1993), assessed by an intraclass correlation.

Coach effectiveness. In addition, overall effectiveness of the coach was rated on a 5-point Likert scale: “Overall, how would you rate this coach?” Respondents could answer from *very poor* (1) to *very good* (5). Coaches and supervisors completed evaluations with the attributional survey and overall effectiveness questions at the end of the evaluation tool.

Procedures

Permission was obtained to use videotapes of coaching sessions by nine students in an undergraduate coaching class, who, as part of their course, were filmed for a practice session to be evaluated by their professor. The students coached recreation level youth football ($n = 5$) and soccer ($n = 4$) teams which ranged in competition level from under six to under fourteen. Consistent with previous research (Ambady and Rosenthal, 1993), three 10 second silent video clips were used from each coach’s session from the beginning, middle, and end; the clips feature the coach alone, consistent with previous research to control for the effects of interaction effects in the environment of the target (Ambady & Rosenthal, 1993).

All of the coaches’ clips were arranged in one videotape in a randomized Latin-square design (Ambady & Rosenthal, 1992). The final tape consisted of 27 clips: 3 clips for each of the 9 coaches.

Each coach rated him/herself on the attribution scale and effectiveness item. Supervisors completed the attribution scale and overall effectiveness item on each coach as part of their formal evaluation of the coach. Evaluations were delivered by the supervisors to the professor and picked up by the researcher.

Raters completed a demographic questionnaire and observed the video of the twenty seven 10-second video clips. Following each clip, raters completed the attributional scale and overall effectiveness question. End-of-the season evaluations by the recreation department supervisors, as well as self-evaluations were used for comparison with the raters' scores on each of the 14 attributes.

Data Analysis

Given that each naïve rater rated each of nine coaches on three occasions, a within-rater mean across three occasions was computed for each coach for each attribute as well as effectiveness. To create an individual difference variable representing target accuracy, 206 correlations between each rater's mean effectiveness scores and supervisor effectiveness scores ($df = 7$) were calculated. To create an individual difference variable representing consensus, 206 correlations between each rater's mean scores and supervisor scores ($df = 7$) were calculated for each attribute. To create an individual difference variable representing self-other agreement, 206 correlations between each rater's mean scores and self scores ($df = 7$) were calculated for each attribute and for effectiveness.

Inferential statistics were utilized to examine moderators of target accuracy, consensus, and self-other agreement. Means were compared using independent sample t-tests for gender comparisons and one-way ANOVAs for comparisons between races and sport participation groups. Post hoc comparisons using Fisher's LSD were conducted on any significant results ascertained from ANOVAs ($p < .01$).

Data Analysis

Individual correlations between each naïve rater's score on effectiveness and the supervisor's score on effectiveness for each coach were calculated and a mean consensus

score was obtained. This provided an individual difference variable representing accuracy accuracy.

Individual correlations between each naïve rater's attributional ratings across nine coaches observed and the supervisors' attributional ratings of these coaches were calculated and a mean correlation was determined to provide an individual difference variable representing consensus. Individual correlations between each naïve rater's attributional ratings across nine coaches observed and the actual coach were calculated and a mean correlation was determined to provide an individual difference variable representing self-other agreement.

Means were compared using independent sample t-tests for gender comparisons and one-way ANOVAs for comparisons between races and sport participation groups. Post hoc comparisons using Fisher's LSD were conducted on any significant results ascertained from ANOVAs: ($p < .01$).

CHAPTER 3

RESULTS

The mean correlations between the naïve raters' effectiveness ratings and the supervisors' effectiveness ratings were calculated to estimate target accuracy of the thin slice judgments by the naïve raters. See Table 1 for results.

Table 1

Descriptive Statistics

	<i>M</i>	<i>SD</i>	Skewness (<i>SE</i> = 0.17)	Kurtosis (<i>SE</i> = 0.34)	<i>M</i>	<i>SD</i>	Skewness (<i>SE</i> = 0.17)	Kurtosis (<i>SE</i> = 0.34)
	Target Accuracy							
Effectiveness	-.27	0.25	0.65	0.69				
Attribute	Consensus				Self-Other Agreement			
Acceptance	-.33	0.28	0.65	0.65	.03	0.30	-0.57	0.33
Active	-.16	0.25	0.10	-0.44	.16	0.28	-0.08	0.30
Attentive	.23	0.27	-0.69	0.91	.11	0.28	-0.20	0.03
Competent	-.15	0.23	0.69	1.40	.19	0.28	-0.15	-0.19
Confidence	.15	0.25	-0.07	0.18	-.05	0.28	0.23	-0.12
Dominance	-.11	0.24	0.30	1.10	.27	0.25	-0.90	-0.05
Empathic	-.17	0.28	0.45	0.56	.42	0.32	-1.20	0.60
Enthusiastic	.45	0.24	-0.99	1.50	-.11	0.30	0.64	1.80
Honesty	-.07	0.27	0.25	-0.10	-.08	0.26	0.33	0.42
Likeability	.20	0.23	-0.22	0.21	.01	0.29	0.48	0.02
Optimistic	.00	0.23	0.13	0.15	.18	0.28	-0.59	0.43
Professional	-.09	0.25	0.02	-0.35	.22	0.27	-0.42	0.10
Supportive	-.17	0.25	0.35	0.11	.01	0.27	0.00	0.10
Warm	-.13	0.28	0.16	-0.10	-.09	0.29	0.23	-0.08

The mean correlations between the naïve raters' ratings on each of the fourteen attributes with the supervisors' ratings on each of the fourteen attributes were calculated to estimate consensus (see table 1).

See Table 1 for results regarding self-other agreement. Independent samples t-tests were run based off of means generated on male and female raters to determine differences between the two groups on accuracy.

There were no differences found on accuracy between groups (see Table 2). Independent samples t-tests run on differences on consensus between genders found significant differences ($p < .01$) on one of the fourteen variables: likeability. Female raters were higher on means consensus than male raters on likeability (see Table 2).

Table 2

Descriptive Statistics for Target Accuracy and Consensus Differentiated by Gender

Attribute	Males		Females	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Effectiveness	-.28	0.28	-.26	0.23
Acceptance	-.33	0.30	-.33	0.26
Active	-.18	0.23	-.14	0.25
Attentive	.20	0.30	.25	0.24
Competent	-.14	0.24	-.16	0.23
Confidence	.24	0.24	.15	0.25
Dominance	-.12	0.25	-.11	0.23
Empathic	-.18	0.33	-.17	0.24
Enthusiastic	.40	0.24	.48	0.23
Honesty	-.08	0.30	-.05	0.24
Likeability*	.14	0.23	.25	0.22
Optimistic	-.04	0.25	.03	0.22
Professional	-.13	0.26	-.07	0.24
Supportive	-.18	0.28	-.15	0.22
Warm	-.13	0.30	-.13	0.27

*significant at $p < .01$

Due to the small sample size of Hispanic, Asian, and Other, they were not included in analyses on race differences. Independent samples t-tests run on differences between Caucasian and African-American raters found no significant differences on accuracy or consensus ($p > .01$) (see Table 3).

Table 3

Descriptive Statistics for Target Accuracy and Consensus Differentiated by Race

Attribute	African-Americans		Caucasians	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Effectiveness	-.26	0.25	-.30	0.25
Acceptance	-.31	0.29	-.39	0.24
Active	-.31	0.29	-.39	0.24
Attentive	.24	0.28	.20	0.22
Competent	-.15	0.22	-.13	0.26
Confidence	.16	0.24	.11	0.26
Dominance	-.09	0.24	-.17	0.18
Empathic	-.17	0.28	-.19	0.29
Enthusiastic	.45	0.25	.42	0.23
Honesty	-.06	0.28	-.09	0.23
Likeability	.19	0.23	.26	0.24
Optimistic	-.02	0.23	.06	0.23
Professional	-.10	0.26	-.07	0.22
Supportive	-.18	0.25	-.15	0.25
Warm	.14	0.28	-.11	0.27

In addition a one-way ANOVA showed no significant differences between levels of sport participation on accuracy ($p > .01$) (see Table 4). However, there were significant differences ($p < .01$) between level of sport participation groups on consensus on two of the fourteen variables: competence, and confidence (see Table 4). Fisher's LSD post hoc

tests indicated that naïve raters who's highest level of competition was at the college level showed significantly more consensus with supervisor ratings on competence than all other categories of level of sport participation raters. College raters also showed significantly more consensus with supervisor ratings on confidence than two other sport participation groups: no participation and varsity/elite participation.

Table 4

Analysis of Variance for Attributes between Levels of Sport Participation

Groups

Attribute	df	F	<i>p</i>
Acceptance	4	0.85	0.50
Active	4	0.29	0.89
Attentive	4	0.96	0.43
Competent*	4	3.57	0.01
Confidence*	4	3.67	0.01
Dominance	4	0.31	0.87
Empathic	4	0.32	0.86
Enthusiastic	4	3.22	0.01
Honesty	4	0.70	0.59
Likeability	4	1.41	0.23
Optimistic	4	0.94	0.45
Professional	4	0.71	0.59
Supportive	4	1.51	0.20
Warm	4	1.45	0.22

*significant at $p < .01$

CHAPTER 4

DISCUSSION

There were several constructs of accuracy measured in this study. The first research question examined the target accuracy of the naïve raters. Due to the lack of correlation between the naïve raters' judgments and the supervisors' evaluations, the naïve raters as a group were not accurate in their assessments of coaching effectiveness. There are several explanations why this may have occurred. The nine coaches varied across two sports and four age levels. They were not observed directly with the athletes so differences in coaching behaviors due to varying age and sport contexts may have caused some of the variability. Thin-slice judgments in the sport context may have more variables that need to be controlled for than thin-slicing in classroom settings or social settings that have been previously examined. Modeling the Ambady and Rosenthal (1992) study the coaches were presented on muted video clips without athletes present. Ambady and Rosenthal presented teachers alone in the clips they showed to naïve raters to control for biases from reactions from students being taught. The coaching context requires adaptations to lessons as well as more frequent feedback. There may be a need for more frequent transactions whereas teaching may include more directive communication. Observations of a coach may require this interaction to accurately assess coaching effectiveness. The design of this study did not allow naïve raters to observe direct interactions between the coach and players.

Another explanation to support the complexity of the sport context is the individual differences in perceptions of effective coaches. Previous research found a negative correlation between body size and perceptions of coaching effectiveness by

female gymnasts, while no correlation was found for soccer players or basketball players (Cumming et al., 2005). This study did not survey for particular sport participation so variation may be due mainly to perceptions of coaching effectiveness in a particular sport. Other research suggests that the personality of the athlete can effect coaching evaluations. Williams et al. (2003) found that athletes with higher anxiety and lower self-confidence rated effectiveness of coaches more negatively. This study did not look at the personality makeup of the raters to determine if those attributes moderate accuracy. Previous research also suggests that mood state can affect evaluations (Ambady & Gray, 2002). Recent research shows that mood state of customers can effect evaluation of salespeople (Puccinelli, 2006). When customers were in a bad mood and the salesperson was perceived as happy the customer rated the salesperson negatively. Ambady and Gray (2002) found that negative mood states affected accuracy of social perceptions.

Another possible explanation why there was not a relationship between naïve raters and head coaches on coaching effectiveness is the lack of congruency between the present circumstances of the raters and the environment of the target. The targets were coaching at the recreation level and the raters were college students. If they had participated in recreation level athletics they were many years removed from the situation. Much of the previous research on thin-slicing has used blind raters who are within the context being evaluated. One example is Ambady and Rosenthal's (1992) study on teacher effectiveness. The naïve raters were college students and they were rating college instructors and their judgments were compared to other student evaluations. This current study used college aged naïve raters who evaluated other college student coaches in a youth sport context. Other studies look at social contexts that

most people are familiar with on a current basis (Ambady, Hallahan, & Rosenthal, 1995). It may be useful to preface the thin-slicing with the context being rated. The naïve raters were not aware they were judging recreation level coaches. It may have been more useful to use parents of children who are in the recreation level context.

Consensus between naïve raters and experts on attributes was not reached on thirteen of the fourteen attributes. Consensus was defined within this study as the agreement between the naïve raters and the expert on personality attributes. Overall significance was not reached on thirteen of the fourteen attributes. Overall consensus was not reached on thirteen of the fourteen attributes. Considering how many correlations were measured, it can be expected that one could reach significance solely by chance. Kenny (1994) defines consensus as the agreement between two raters. This research treats the naïve raters as one and the expert as the second rater. Consensus operationalized this way shows if naïve raters view a target similarly to a person who has greater knowledge of the target.

This approach has limitations because the naïve raters are compared with only one knowledgeable rater. Previous research suggests that there is greater accuracy in judgments of a target when there are two or more evaluations from people who know the target (Kolar, Funder, & Colvin, 1996). Consensus may have been higher if more than one judgment by knowledgeable others could have been averaged to determine consensus.

Consensus in Ambady and Rosenthal's (1993) research was operationalized by intracorrelations of naïve raters' judgments of attributes which were placed in a 15 X 15 matrix and subjected to a principle components analysis. It is possible that consensus

between naïve raters was reached in this study, which means they could have viewed the target similarly. This is a research question that should be considered for future research.

In regards to consensus, there was a moderate relationship between naïve raters and supervisors on the attribute enthusiastic. Previous research on the Norman and Goldberg's (1966) Big Five and zero acquaintance research found consensus on the extraversion factor of the Big Five (Funder & Dobroth, 1987; Kenny, Albright, Malloy, & Kashy, 1994; Park & Judd, 1989). Characteristics suggested by the extraversion category include sociable and energetic. It is possible that enthusiastic may be very similar to, or an expression of, extraversion. It could be easier to observe than the other traits. Researchers (Kenny, Albright, Malloy, & Kashy, 1994) suggest that extraversion is processed very quickly. John and Robins (1993) suggest that the observability and evaluativeness of the attributes can contribute to accuracy and agreement between raters. The more neutral (less evaluative) and observable an attribute is the greater the agreement between raters is about the target. For example talkativeness is observable and neutral, while arrogance could be viewed as negative and more difficult to observe. Most of the fourteen attributes in this study were positively charged and difficult to directly observe: accepting, attentive, competent, confident, dominant, empathic, enthusiastic, honest, likable, optimistic, professional, supportive, and warm.

Little research has examined thin-slicing in the sport context. Potentially personal biases of raters could affect judgments of coaches' attributes. Kenny (1994) explains that "personal stereotypes", such as whether a rater subscribes to a widely held view. An example would be "all professors are absent-minded", which can be reflected in judgments, and does not necessarily change with increasing acquaintance. Current

research shows that stereotypes are based on more than gender or race. Kenny (1994) explains that appearance cues and nonverbal behaviors are associated with different personality traits.

There was not self-other agreement in this study between the naïve raters' judgments and the coaches self judgments of personality attributes. Previous perception research found that self judgments were less accurate when assessing behavior than others (Kolar, Funder, & Colvin, 1996; Spain, Eaton, Funder, 2000). Robins and John (1997) suggest that mood affects self judgments as well as the need to protect self-esteem. The coaches in this study were undergraduate college students with no previous coaching experience. Their own perceptions about their coaching may have entered into the answering of the survey questions. Coaching literature has found that coaches are unaware of how they present themselves and behave while coaching (Smith & Smoll, 1997; Smith, Smoll, & Curtis, 1978; Wandzilak, Ansorge, & Potter, 1988). It is possible that the coaches in this study are similar and unaware of their behaviors.

This study supports the research literature in which no significant differences were found between gender and target accuracy. This supports an earlier meta-analysis by Ambady and Rosenthal (1992) that examined numerous studies and concluded that overall gender did not affect thin slicing or zero acquaintance judgments. It has been suggested that women are better judges of nonverbal behavior (Hall, 1984). Rosenthal and DePaulo (1979) found that women are better judges when the information is presented in more controllable channels. Speech is considered the most controllable channel, while the voice is considered the least controllable (Brown, 1986). This study

did not involve an auditory component so potential differences in gender may not have arisen because of the channels for the cues of nonverbal behavior.

There was a significant difference between male and female naïve raters on one of the fourteen attributes. The most only significant difference ($p < .01$) was for likeability attribute. Female raters were closer to consensus with supervisors than male raters. This may pertain to the different expectations by gender on participation in sport. Previous studies have shown that females emphasize friendship and social interaction over competition and achievement than males do (Alderman, 1988; Gill, Gross, & Huddleston, 1983; Gould, Feltz, & Weiss, 1986; Petrie, 1971). Dubois (1990) found that the longer youth participate in sport the greater the divergence in values placed on the outcomes by gender, Experienced males place greater importance on outcomes, whereas females consistently place emphasis on social aspects of sport. Potentially female raters in this study may have been more attuned to characteristics that embody the outcomes they desire in a sport setting. The other two attributes in which females differed significantly from males were enthusiastic and optimistic. All three of the differences between variables could be explained by the greater emphasis females place on these attributes and potentially the greater awareness they have of these attributes.

Overall there were no differences between African-Americans and Caucasians on target accuracy or consensus. Little research has examined racial differences in perception of naïve raters. Previous research has found race of target to affect accuracy and consensus (Chiao et al., 2006; Gosling et al., 2002). This research shows that race of raters does not affect target accuracy or consensus. Perhaps the sport context is different due to the length of participation of different races in sport and public acceptance of

different races in sport over other areas in society. Edwards (1973) suggests that lack of opportunities in mainstream society due to discrimination has led a disproportionately high number of blacks to pursue sport. Bledsoe (1973) highlighted the practice in which young blacks pursue sport because of the lack of successful black role models in other areas. Sport is an area that has provided opportunity for those lower on the socioeconomic ladder to gain recognition and money when other avenues were closed off to them. (Coakley, 2001). This can be supported by statistics: blacks make up 77% of the NBA, 64% of the WNBA, and 65% of the NFL, they are only 4.2% of our physicians, 2.7% of our lawyers and 2.2% of our civil engineers (Bureau of Labor Statistics, 1997). In NCAA Division I athletics blacks comprise 23.5% of student athletes: black males = 29.5% of male athletes; black females = 14.2% of female athletes). Black males comprise 60% of basketball players and 51% of football players and 27% of track athletes, while black females constitute 35% of basketball players and 31% of track athletes (NCAA, 1998).

Perceptions of the race of the coaches may have also played a role in the lack of significant differences between races. There were eight Caucasian coaches and one African American coach. Statistics show a disproportionate number of non-Latino white males in coaching positions in the professional leagues and NCAA (Lapchick & Mathews, 2000). “Stacking” theories in sport studies suggest that blacks are placed in positions that require more speed and stamina but less cognitive processes. One result of this is less opportunity to coach for minorities because of the positions they played that required less understanding of the overall game (Coakley, 2001). There is a pattern found

in professional sports and college sports of a disproportionately high number of blacks playing on teams coached by whites. (Coakley, 2001).

Overall there were no differences among levels of sport participation of raters on consensus of effectiveness. There was no correlation with the criterion variable between sport participation groups. Eight of the nine coaches were rated by supervisors as a four or a five out of five on effectiveness. The ninth coach was rated a three. Naïve raters overall rated coaches less effective than the supervisors. This could be a function of expectations of effective coaches at different levels. These coaches are fulfilling a requirement of an undergraduate coaching course which meets 3 hours a week. These coaches may experience more instruction which affects their ratings by supervisors.

While there were not significant differences in most of the attributional categories, there were significant differences on two of the fourteen attributes among levels of sport participation of raters. The higher the level of sport participation the greater the consensus with the expert judge on the competence attribute: the raters with college participation were significantly different than raters with varsity/elite experience, junior varsity experience, recreation level experience, and no sport experience. The college level athletes had greater consensus than all the other groups. One explanation could be the greater participation of these raters in sport and their level of attunement to competence of coaches. These raters possibly had a greater exposure to a number of coaches and are more sensitive to competence. Millard (1996) posits that the higher level an athlete pursues the greater the need for winning and the greater the need for technical instruction from a coach. She found that coaches who provided more instruction based feedback were perceived as more competent. High-experience coaches are noted to

provide more technical feedback and less general encouragement than low-experience coaches (Sherman and Hassan, 1986). This difference could also account for the awareness of competence of the college level raters.

The college level raters were also significantly different than varsity/elite athletes and recreation level athletes on confidence. The college level raters showed more consensus with supervisors' ratings. They could also be attuned to the confidence level of coaches. Research shows that male coaches are generally more confident in abilities than female coaches (Millard, 1996). This study used eight male coaches and one female coach. College level raters due to length involved in sport may be more attuned to the confidence level of a coach.

Researchers attempt to define the moderators surrounding the rater, the channel, the judgments, and the target that could affect accuracy. It is also valuable to learn in what scenarios judgments are not accurate. Evans (1984) notes that it is more important to know in what contexts people do not make good decisions. Previous research suggests that the degree to which a judge cares about the judgment he or she is making can affect the accuracy and consensus (Flink & Park, 1991; Funder, 1995). The environment observed may have also affected consensus on personality judgments. Previous research found that less structured situations yield greater correlations on personality (Funder & Colvin, 1991; Snyder & Ickes, 1985). This research involved judgments of targets in a classroom setting observing video clips instead of directly observing the targets in the sport environment.

This research is promising because it is the first to examine thin-slicing in the sport environment. It suggests that the sport context may have more variables to control

for when doing zero acquaintance research. Future research should attempt to control variables and look at particular sports and use naïve raters who have experienced that sport. Future research could also examine zero acquaintance situations at different levels, like the collegiate or elite level. Looking at moderators of consensus based on the demographics of the coach, like gender and race would be valuable. Qualitative studies could further understand personal biases that underscore perceivers' views of effective coaches, whether gender, sport level and type, or race could affect that.

REFERENCES

- Alderman, R. B. (1988). Strategies for Motivating Young Athletes. In W. F. Straub (Ed.) *Sport psychology: An analysis of athlete behavior* (pp. 4961). Ithaca, NY: Movement.
- Abromovitch, R. (1977). Children's recognition of situational aspects of facial expressions. *Child Development, 48*, 459-463.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.
- Allport, G. W., & Vernon, P. E. (1933). *Studies in expressive movement*. New York: Haffner.
- Ambady, N., Conroy, M., Mullins, J., & Tobia, A. (2001). *Friends, lovers, and strangers: Judging dyadic relationships from thin slices*. Manuscript submitted for publication.
- Ambady, N., & Gray, H. M. (2002). On being sad and mistaken: Mood effects on the accuracy of thin-slice judgments. *Journal of Personality and Social Psychology, 83*, 947-961.
- Ambady, N., Hallahan, M., Rosenthal, R. (1995). On judging and being judged accurately in zero-acquaintance situations. *Journal of Personality and Social Psychology, 69*, 518-529.
- Ambady, N. & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*, 256-274.

- Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, *64*, 431-441.
- Archer, D., & Akert, R. M. (1977). Words and everything else: Verbal and nonverbal cues in social interpretation. *Journal of Personality and Social Psychology*, *35*, 443-449.
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, (Vol.6, pp. 1-62). New York: Academic Press.
- Bledsoe, T. (1973). Black dominance in sports: Strictly from hunger. *The Progressive*, *37*, 16-19.
- Bloom, B. (1985). *Developing talent in young people*. New York: Ballantines.
- Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, *62*, 645-657.
- Brown, R. (1986). *Social psychology: The second edition*. New York: Free Press.
- Bureau of Labor Statistics (1997). Retrieved June 6, 2006 from <ftp://ftp.bls.gov/pub/special.requests/lf/aat11.txt>
- Chiao, J. Y., Heck, H. E., Nakayama, K., & Ambady, N. (2006). Priming race in biracial observers affects visual search for black and white faces. *Psychological Science*, *17*, 387-392.
- Coakley, J. (2001). *Sport in society: Issues and controversy*. New York: McGraw-Hill.
- Conroy, D. E., Coatsworth, J. D. (2006). *Coach training as a strategy for promoting youth social development*, *20*, 124-144.

- Csikszentmihalyi, M., & McCormack, J. (1986). The influence of teachers. *Phi Delta Kappan*, 66, 415-419.
- Cumming, S. P., Eisenmann, J. C., Smoll, F. L., Smith, R.E., Malina, R. M.(2005). Body size and perceptions of coaching behaviors by adolescent female athletes. *Psychology of Sport and Exercise*, 6, 693-705.
- Dubois, P. (1990). Gender differences in value orientation towards sports: A longitudinal analysis. *Journal of Sport Behavior*, 13, 3-14.
- Edwards, H. (1973). *Sociology of sport*. Dorsey Press: Homewood, IL.
- Eyal, N., Bar-Elis, M., Tenenbaum, G., & Pie, J. S. (1995). Manipulated outcome expectations and competitive performance in motor tasks with gradually increasing difficulty. *The Sport Psychologist*, 9, 188-200.
- Evans, J. St. B. T. (1984). In defense of the citation bias in the judgment literature. *American Psychologist*, 39, 1500-1501.
- Feltz, D. L., & Riessinger, C. A. (1990). Effects of in vivo emotive imagery and performance feedback on self-efficacy and muscular endurance. *Journal of Sport and Exercise Psychology*, 12, 132-143.
- Flink, C., & Park B. (1991). Increasing consensus in trait judgments through outcome dependency. *Journal of Experimental Social Psychology*, 27, 453-467.
- Flora, C. (2004). The once-over can you trust first impressions. *Psychology Today*, 37, 60-65.
- Fortunato, V. J., & Mincy, M. D. (2003). The interactive effects of dispositional affectivity, sex, and a positive mood induction of student evaluations of teachers. *Journal of Applied Sport Psychology*, 33, 1945-1968.

- Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science, 2*, 31-39.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*, 652-670.
- Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology, 60*, 773-794.
- Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52*, 409-418.
- Gill D. L., Gross, J. B., Huddleston, S. (1983). Participation motivation in youth sports. *International Journal of Sport Psychology, 14*, 1-14.
- Gladwell, M. (2005). *Blink*. New York: Little, Brown & Company.
- Gould, D., Feltz, D., Weiss, M. (1986). Motives for participating in competitive youth swimming. *International Journal of Sport Psychology, 16*, 126-140.
- Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology, 82*, 379-398.
- Grahe, J. E., & Bernieri, F. J. (1999). The importance of nonverbal cues in judging rapport. *Journal of Nonverbal Behavior, 23*, 253-269.
- Greelees, I., Bradley, A., Holder, T., & Thelwell, R. (2005). The impact of opponents' non-verbal behavior on the first impressions and outcome expectations of table-tennis players. *Psychology of Sport & Exercise, 6*, 103-115.

- Hall, J. A. (1984). *Nonverbal sex differences: Communication accuracy and expressive style*. Baltimore: Johns Hopkins University Press.
- Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal of Personality*, 8, 149-162.
- John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, 61, 521-551.
- Jones, E. E., & Nisbett, R. E. (1971). *The actor and the observer: Divergent perceptions of the causes of behavior*. Morristown, NJ: General Learning Press.
- Kahlbaugh, P. E., & Haviland, J. M. (1994). Nonverbal communication between parents and adolescents: A study of approach and avoidance behaviors. *Journal of Nonverbal Behavior*, 18, 91-113.
- Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York, New York: Guilford Publications. Retrieved March 16, 2006 from <http://davidakenny.net/ip/assim.htm>
- Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus and interpersonal perception: Acquaintanceship and the Big Five. *Psychological Bulletin*, 116, 245-258.
- Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L. (1992). Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology*, 62, 88-97.

- Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of personality judgments by the self and knowledgeable others. *Journal of Personality, 64*, 311-337.
- Kruglanski, A. W. (1989). The psychology of being “right”: The problem of accuracy in social perception and cognition. *Psychological Bulletin, 106*, 395-409.
- Lapchick, R., & Mathews, K. (2000). *1998 Racial and gender Report Card*. The Center for the Study of Sport in Society: Boston, MA.
- Millard, L. (1996). Differences in coaching behaviors of male and female high school soccer coaches. *Journal of Sport Behavior, 19*, 19-32.
- McGinnis, A. L. (1985). *Bring out the best in people*. Minneapolis: Augsburg Publishing House.
- NCAA (1998). *Race demographics of NCAA member institutions' athletics personnel*. The National Collegiate Athletics Association: Overland Park, Kansas.
- Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality and Social Psychology, 4*, 681-691.
- Park, B., & Judd, C. M. (1989). Agreement of initial impressions: Differences due to perceivers, trait dimensions, and target behaviors. *Journal of Personality and Social Psychology, 56*, 493-505.
- Petrie, B. M. (1971). Achievement orientations in adolescent attitudes toward play. *International Review of Sport Sociology, 6*, 89-101.
- Puccinelli, N. M. (2006). Putting Your Best Face Forward: The Impact of Customer Mood on Salesperson Evaluation. *Journal of Consumer Psychology, 16*, 156-162.

- Robins, R. W., & John, O. P. (1997). The quest for self-insight: Theory and research on accuracy and bias in self-perception. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 649-679). San Diego, CA: Academic Press.
- Rosenthal, R. (1987). *Judgment studies: Design, analysis, and meta-analysis*. New York: Cambridge University Press.
- Rosenthal, R., & DePaulo, B. M. (1979). Sex differences in accommodation in nonverbal communication. In R. Rosenthal (Ed.), *Skill in nonverbal communication*. Cambridge, MA: Oelgeschlager, Gunn, & Hain.
- Sherman, M. A., & Hassan, J. S. (1986). In M. Pieron, & G. Graham (Eds.), *The 1984 Olympic scientific congress proceedings* (Vol. 6) (pp. 103-108). Champaign, IL: Human Kinetics.
- Slavin, R. E. (1980). Cooperative learning. *Review of Educational Research*, 50, 312-342.
- Smith, R. E., & Smoll, F. L. (1997). Coaching the coaches: Youth sports as a scientific and applied behavioral setting. *Current Directions in Psychological Science*, 6, 16-21.
- Smith, R. E., & Smoll, F. L. (1990). Self-esteem and children's reaction to youth sport behaviors: A field study of self-enhancement processes. *Developmental Psychology*, 26, 987-993.
- Smith, R. E., Smoll, F. L., & Curtis, B. (1978). Coaching behaviors in Little League baseball. In F.L. Smoll & R.E. Smith (Eds.), *Psychological perspectives in youth sports* (pp. 173-201). Washington, DC: Hemisphere.

- Smith, R. E., Smoll, F. L., & Curtis, B. (1979). Coach Effectiveness Training: A cognitive-behavioral approach to enhancing relationship skills in youth sport coaches. *Journal of Sport Psychology, 1*, 59-75.
- Smoll, F. L., Smith, R. E., Barnett, N. P., & Everett, J. J. (1993). Enhancement of children's self-esteem through social support training for youth sport coaches. *Journal of Applied Psychology, 78*, 602-610.
- Snyder, M., & Ickes, W. (1985). Personality and social behavior. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 2, pp. 883-948). New York: Random House.
- Spain, J.S., Eaton, L.G, & Funder, D.C. (2000). Perspectives on personality: The relative accuracy of self versus others for the prediction of emotion and behavior. *Journal of Personality, 68*, 837-867.
- Tickle-Degnan, L., & Rosenthal, R. (1987). Group rapport and nonverbal behavior. In C. Hendrick (Ed.), *Group processes and intergroup relations. Review of personality and social psychology* (Vol. 9, pp. 113-136). Beverly Hills, CA: Sage.
- Tickle-Degnan, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry, 4*, 285-293.
- Vallee, C. N., & Bloom, G. A. (2005). Building a successful university program: Key and common elements of expert coaches. *Journal of Applied Sport Psychology, 17*, 179-196.
- Vrij, A., Dragt, A., & Koppelaar, L. (1992). Interviews with ethnic interviewee: Non-verbal communication errors in impression formation. *Journal of Community & Applied Social Psychology, 2*, 199-208.

- Wandzilak, T., Ansorge, C.J., & Potter, G. (1988). Comparison between selected practice and game behaviors of youth sport soccer coaches. *Journal of Sport Behavior*, 2, 78-88.
- Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology*, 57, 120-128.
- Weinberg, T., Gould, D., & Jackson, A. (1979). Expectations and performance: An empirical test of Bandura's self-efficacy theory. *Journal of Sport Psychology*, 3, 345-354.
- Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.
- Williams, J. M., Jerome, G. J., Kenow, L. J., Rogers, T., Sartain, T. A., & Darland, G. (2003). Factor structure of the coaching behavior questionnaire and its relationship to athlete variables. *Sport Psychologist*, 17, 16-35.
- Wooden, J. R. (1980). *Practical modern basketball*. New York: John Wiley and Sons.

APPENDICES

APPENDIX A
RESEARCH QUESTIONS, LIMITATIONS, DELIMITATIONS, ASSUMPTIONS,
DEFINITIONS OF TERMS

Research Questions

1. What is the naïve raters' accuracy in their assessment of coaching effectiveness?
2. What is the consensus between naïve raters (given only thin slices) and experts on each attribute?
3. What is the self-other agreement between naïve raters and coaches on each attribute?
4. Is there a significant difference in accuracy between male and female raters?
5. Is there a significant difference in consensus between male and female raters?
6. Is there a significant difference in accuracy between races of raters?
7. Is there a significant difference in consensus between races of raters?
8. Is there a significant difference in accuracy between raters' level of sport participation?
9. Is there a significant difference in consensus between raters' level of sport participation?

Limitations

1. Raters will not be randomly selected.
2. Raters will be in a large room together which may cause bias in ratings.
3. Raters will not be compensated in any way which may affect their willingness to participate honestly.
4. Raters may possess internal biases that will affect their ratings.
5. Raters may not have enough time to answer each survey due to time constraints in the classroom.
6. Participants will be videotaped for only one practice session.

7. Participants may act differently while being videotaped during practice.
8. There is no validity or reliability to the evaluation questionnaire used for the participants.
9. There is no validity or reliability on the attributional list which is used for the coaches and raters.
10. Supervisors may evaluate with answers that are socially desirable responses, not personally accurate responses.
11. Participants may not understand some of the questions.
12. The researcher can not control the interest level of the participants.
13. Attributes being rated are not neutral in judgment and may not be highly observable; high observability and low evaluativeness of attributes makes judgments easier for raters (John & Robins, 1993).

Delimitations

1. The sample only involves undergraduate coaching students from one southeastern university.
2. The sample only uses recreation level coaches.
3. There is a narrow age range of participants.
4. Only quantitative results are available from the questionnaires.

Assumptions

1. Participants will give honest responses to the demographics survey, the effectiveness rating, and the attributional scores.
2. Supervisors will give honest responses to the evaluation, the effectiveness rating, and the attributional scores.

3. Participants will understand the verbal and written instructions given to them by the researcher.
4. Participants will complete all research instruments.

Definitions

Zero Acquaintance- a condition in which perceivers make judgments about targets, but the two have never engaged in social interaction (Kenny et al., 1992).

Thin-slicing as part of zero acquaintance research-defined by Rosenthal and Ambady as a brief excerpt of expressive behavior sampled from the behavioral stream (Ambady & Gray, 2002).

Naïve raters- perceivers in zero acquaintance situations where the perceiver and target have never met and a judgment is made about the target.

Target accuracy in thin-slicing-the correspondence between participants' judgments of a target individual and well-defined external criterion (Rosenthal & Ambady, 1992; Ambady & Hall, 2002). Target accuracy was operationalized by the mean consensus between naïve raters' judgments of overall effectiveness with the ecologically valid criterion (supervisors' evaluation of overall effectiveness) (Ambady & Hall, 2002).

Consensus- Accuracy defined by the degree to which two perceivers rate a target the same way (Kenny, 1998). Consensus was operationalized by the mean consensus between naïve raters' judgments of attributes with the supervisors' judgments of attributes.

Coach effectiveness-the degree to which an individual fulfills his/her coaching duties

Attributes-observable behavioral characteristics

Self-other agreement- Accuracy defined by the degree to which a person rates him or herself as do others (Kenny, 1998). Self-other agreement was operationalized by the mean consensus between naïve raters' judgments of attributes with the supervisors' judgments of attributes.

APPENDIX B
EXTENDED REVIEW OF LITERATURE, EXTENDED REVIEW OF
LITERATURE REFERENCES

Thin-Slicing Background

“This remarkable capacity we possess to understand something of the character of another person, to form a conception of him as a human being ... with particular characteristics forming a distinct individuality is a precondition of social life” (Asch, 1946, p. 258).

It is suggested that we communicate our interpersonal expectancies and biases through nonverbal cues, which are neither encoded nor decoded at a conscious level of awareness (Chaikin, Sigler, & Derlega, 1974; Christensen & Rosenthal, 1982; Harris & Rosenthal, 1985; Rosenthal, 1966; Rosenthal & Rubin, 1978; Snyder, Tanke, & Berscheid, 1977; Word, Zanna, & Cooper, 1974). It is suggested that nonverbal behaviors are more indicative of true meaning when conflict exists between verbal and nonverbal behaviors. Researchers found that judges who expected a defendant to be guilty revealed their expectation in their nonverbal behavior, not in their verbal behavior (Blanck, Rosenthal, Vannelli, & Lee, 1986). Married couples who were distressed but trying to act happy were distinguishable from happy couples based on nonverbal behaviors (Vincent, Friedman, Nugent, & Messerly, 1979). Expressive behaviors are expected to be more controllable in order to portray socially desirable responses (Goffman, 1959, 1971); less controllable behavioral channels allow true feelings to “leak” out (Ekman & Friesen, 1969). Speech is the most controllable channel, followed by the face, the body, and the voice (Brown, 1986; Rosenthal & DePaulo, 1979).

Research on consensus between raters and a target determined that consensus among judges exists but does not explain the bulk of the variance; consensus is highest for extraversion; some degree of consensus is reached even when the judges have

virtually no interaction with the targets (Funder & Colvin, 1988; John & Robins, 1993; Norman and Goldberg, 1966; Watson, 1989). Researchers determined in his analysis of 32 previous social relations studies that longer acquaintanceship does not necessarily lead to higher levels of consensus (Kenny, Albright, Malloy, & Kashy, 1994).

Research on Self-Ratings

Research in the area of others' versus self-judgments has focused on the comparison of accuracy and the particular circumstances in which they are effective. Previous research concerning the accuracy of naïve raters' judgments found a correlation with self-judgments (Funder & West, 1993). The premise was that self-reports were more accurate measures of personality than knowledgeable others' judgments (Wiggins, 1973). They are the standard by which other measures are gauged (Paunonen, 1989). Kolar, Funder, and Colvin (1996) challenged the assumption by comparing the predictive validity of self and others' reports on behaviors. They hypothesized that others' predictions would be more accurate because of the focus on the target's behavior versus the others' focus on the environment. They found that others' predictive abilities on behaviors were stronger than self-judgments when two raters were involved. Spain, Eaton, and Funder (2000) expanded the research by determining the accuracy of others' and self-judgments in rating behaviors and moods of target. The research supported Kolar, Funder, and Colvin in the accuracy of others' ratings over self-reports on behaviors. On the other hand, they found in regard to emotional experience, that self-reports were more accurate.

Regarding self-other agreement, Kenny et al. (1994) found when evaluating the Big Five traits, that agreement was highest for extraversion. John and Robins (1993)

elucidated the relationship further by breaking down trait properties into observability and evaluativeness. They define observability by the degree to which behaviors are relevant to the trait can be easily observed. They define evaluativeness by the degree to which a trait is relatively neutral. They found that interjudge and self-other agreement are highest when the trait is observable and relatively neutral in evaluation. They contend that Kenny's previous findings on self-other agreement on the Norman and Goldberg's (1966) Big Five personality dimensions support their findings. The greatest agreement on extraversion is explained by the relative observability and neutrality of each of the personality characteristics comprising that domain. This supports earlier research (Bem, 1972) that suggests the process on self and other perception is altered when the traits being observed are evaluative. Robins and John (1997) suggest four factors that could affect self-judgments: including, people protect their self-esteem when making evaluative judgments and mood exerts stronger effects on self-judgments.

Accuracy in interpersonal perception

The evaluation of accuracy in interpersonal perception is defined simply as a correspondence between a judgment and a criterion (Brown, 1965; Kruglanski, 1989). Due to the difficulty in identifying objective criteria, Kenny and Albright (1987) identified criterion measures for use in social perception. These measures include self-report measures, third person (expert) judgments, and objective measurements such as physiological variables, judge ratings, and operational criteria such as those used in lie-detection studies. Ambady and Rosenthal (1992) expanded the items identified to include criteria that were ecologically valid and commonly used in everyday decisions about people. Ambady and Rosenthal (1992) explained the reasoning behind using student

evaluations of teachers as ecologically valid due to it being a primary method of making decisions about performance and promotion.

Ambady and Rosenthal (1992) completed a meta-analysis on thin slices of behavior as predictors of interpersonal consequences. The findings indicated that predictions based on less than ½ minute in length did not differ significantly from predictions based on four and five minute observations (Ambady & Rosenthal, 1992). The meta-analysis of 38 previous studies also showed that the behavioral channel (face, speech, the body, tone of voice) was not related to the accuracy of predictions.

Theoretical models in interpersonal perception

Thin slicing is part of the overall body of research of interpersonal perception with social psychology. Warr and Knapper (1968) suggest a framework to explain the person perception process. The initial phase is the entrance of the perceiver (rater or judge) into the social interaction and the attempt of the perceiver to make sense of the interaction. It is through the initial interaction that the perceiver tries to predict the outcome. In turn there are three interdependent responses. First, there is an 'attributive response' which aids in assessing characteristics, traits, states and goals of the target. Second, there is an 'affective response' which explains the emotional responses of the perceiver. Third, there is an 'expectancy response' which explains the expectations the perceiver gains about the target's actions (Warr & Knapper).

Different theoretical models have been used in person perception. The PERSON model is a modification of the weighted-average model (WAM) (Kenny, 1994). WAM has been used to explain why consensus does not increase with increasing acquaintance (Kenny, 1994). It has been used in several studies on interpersonal perception (Borkenau

& Liebler, 1993; Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2003; Chaplin & Panter, 1993; Funder, Kolar, & Blackman, 1995; Malloy, Agatstein, Yaras, & Albright, 1997). The PERSON model allows researchers to develop hypotheses and it will explain results. The model can also be used in conjunction with the social relations model (SRM) (Kenny, 1994). Both focus on variances in interpersonal perception, as opposed to WAM which focuses on correlations. The SRM is a more generalized statistical model that assumes a perceiver's judgment is a function of three components: perceiver, target, and relationship.

A general theoretical model of interpersonal perception called PERSON (personality, error, residual, stereotype, opinion, and norm) supports the framework for research involving a target, a rater and judgments. The model explains the low level of consensus in person perception, the failure to find an increase in consensus as a function of acquaintance, the near perfect test-retest correlations in target effects, the relative validity of other versus self-perception, and the over-confidence effect (Kenny, 2004).

Attributes and teaching effectiveness

Ambady and Rosenthal's (1993) results were consistent with previous research that suggests that certain personality attributes are related to student's ratings of teachers. Attributes correlated with effectiveness, confidence, dominance, warmth and supportiveness resonate with previous research that found achievement-oriented positive behaviors and interpersonally oriented behaviors to be predictors of positive student evaluations (Brophy & Good, 1974; Erdle, Murray, & Rushton, 1985; Rosenshine, 1971; Rosenshine & Furst, 1973, Ryans, 1960).

Research found that teachers rated as effective and extraverted were more nonverbally active and expressive and identified by rapid body movements and smiling (Ambady & Rosenthal, 1993; Kenny, Horner, Kashy, & Chu, 1992). In addition, more sociable and extraverted participants in a zero-acquaintance situation were rated more accurately than those who were less sociable (Ambady, Hallahan, & Rosenthal, 1995; Borkenau & Liebler, 1992; Kenny, Horner, Kashy, & Chu, 1992). Perhaps the accuracy of thin-slicing for teachers is the tendency for more extraverted people to be drawn to a career working with people.

Ambady and Rosenthal (1993) suggest that some stable quality of the teacher was being communicated in the thin-slices of nonverbal behavior. The power of our beliefs and characteristics on our nonverbal behavior has potential strong impacts. Ambady, Koo, Winograd, and Rosenthal (2002) found a link between physical therapists' nonverbal communication and the health outcomes of geriatric patients. Based on a convergence between strangers' ratings of intelligence and the targets' psychometric intelligence, Borkenau and Liebler (1992) suggest that nonverbal behaviors are manifestations of internal traits. They suggest that personality and intelligence are appropriately inferred because of the external manifestation. Early research suggests that personality may be expressed by observable attributes, similar to emotions expressed by facial and vocal manifestations (Scherer, 1978).

Coaching and Teaching Comparisons

Feltz, Chase, Moritz, and Sullivan (1999) posit that athletic coaches are also teachers. Research has shown that self-efficacy is a powerful variable in teacher effectiveness (Ashton & Webb, 1986; Coladarci, 1992; Denhem & Michael, 1981;

Gibson & Dembo, 1984; Guskey & Passaro, 1994; Hoy & Woolfolk, 1993). Feltz and Chase developed the Coaching Efficacy Model on coaching efficacy expecting to find the same relationship. Research showed that coaching efficacy is a powerful variable in coaching effectiveness (Feltz, Chase, Moritz, & Sullivan).

Feltz, Chase, Moritz, and Sullivan (1999) further compared coaches and teachers because both provide instruction, guide the practice of skills, and give feedback to their athletes. In addition, coaches rely on their education, preparation, and experience to be effective. Gould (1987) suggests that coaches are concerned with the learning and performance of their athletes and to be effective, must perform multiple roles.

Expectancy theory research also shows parallels between the teaching and coaching contexts. Rosenthal and Jacobson (1968) suggested that a student will conform to the expectations that a teacher has of him/her. Expectancy theory or self-fulfilling prophecy occurs when a teacher or coach perceives his or her students differently, and provides feedback accordingly. When a teacher has a low-expectancy for a student then he/she will provide less frequent praise (Martinek, 1988; Solomon, et. al, 1996); ask fewer questions (Martinek & Karper, 1984a, 1984b, 1986); give less technical correction (Rejeski, Darracott, & Hutskar, 1979); provide fewer constructive comments (Markland & Martinek, 1988); interact less frequently (Martinek & Johnson, 1979; Sinclair & Vealey, 1989); listen less and criticize more (Martinek & Karper, 1984a, 1984b, 1986). Horn (1984) demonstrated something similar, called the circular effect in the athletic setting. Her research illuminated the initial behavioral differences toward low-expectancy athletes by the coaches, and the subsequent decrease in interaction by the athlete and withdrawal in the athletic setting. Good and Brody (1987) also supported the circular

theory in the classroom. Sinclair and Vealey (1989) also observed that a coach's communication of different expectations are often unconsciously presented as a function of general biases or past instructional experiences with particular students (Sinclair & Vealey, 1989).

Coaching Effectiveness

Coaching effectiveness is the projected outcome of Smith, Smoll, and Curtis' (1979) Coach Effectiveness Training (CET). While personality characteristics can be inferred from the stages, the coach is taught to focus on correcting behaviors (i.e. redefining winning, focus on positive reinforcement, establish social norms, involve the athletes in establish team rules, and receive behavioral feedback for self-monitoring). Measurement of these behaviors can be accomplished through the use of the Coaching Behavior Assessment System (CBAS) (Smith, Smoll, & Hunt, 1977). More recent research has expanded the purpose of the CBAS to include the evaluative reactions of the athletes in determining effectiveness of particular coaching behaviors. The Coaching Behavior Questionnaire (CBQ) also determines attributes of the athletes and their prediction in coach effectiveness (Williams, et al., 2003). An athlete with higher anxiety and lower self-esteem is more likely to negatively evaluate a coach's behaviors.

Kenow and Williams (1992) suggest that given behaviors in a given situation can not be expected to produce consistency in effect on all athletes. Little research has been found that discusses the impact of a coach's personal attributes and the subsequent nonverbal behavior of coaches in conducting what is perceived by athletes to be effective coaching methods.

Chelladurai's multidimensional model of leadership (1980, 1990) predicts several factors, including a coach's personal characteristics (e.g., gender, age, years of experience, psychological characteristics), that influence the coach's actual behaviors. While research supports other relationships in his model (Chelladurai, 1984; Chelladurai & Carron, 1983; Chelladurai, Malloy, Imamura, & Yamaguchi, 1987; Riemer & Chelladurai, 1995; Weiss & Friedrichs, 1986), the relationship between psychological characteristics and coaching behaviors and athlete satisfaction has yet to be explored (Horn, 1992). More recent research has found a link between a coach's burnout level and subsequent coaching behaviors (Price & Weiss, 2000; Vealey, Armstrong, Comar, & Greenleaf, 1998). Burnout is recognized as a psychological syndrome that takes on characteristics like low self-esteem, depersonalization, and emotional exhaustion (Maslach & Jackson, 1981a, 1981b). Price and Weiss reported that coaches with higher burnout were recorded with less training and instruction and social support and as making fewer autocratic and more democratic decisions.

ADDITIONAL REVIEW OF LITERATURE REFERENCES

- Ambady, N., Hallahan, M., Rosenthal, R. (1995). On judging and being judged accurately in zero-acquaintance situations. *Journal of Personality and Social Psychology, 69*, 518-529.
- Ambady, N., Koo, J., Winograd, C. H., & Rosenthal, R. (2002). Physical therapists' nonverbal communication predicts geriatric patients' health outcomes. *Psychology & Aging, 17*, 443-453.
- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*, 256-274.
- Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology, 64*, 431-441.
- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 258-290.
- Ashton, P. T., & Webb, R. B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. New York: Longman.
- Blanck, P. D., Rosenthal, R., Vannelli, M., & Lee, T. D. (1986). Therapists' tone of voice: Descriptive, psychometric, interactional and competence analyses. *Journal of Social and Clinical Psychology, 4*, 154-178.
- Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62*, 645-657.

- Borkenau, P., & Liebler, A. (1993). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality and Social Psychology*, 65, 546-553.
- Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology*, 86, 599-614.
- Brophy, J. E., & Good, T. L. (1974). *Teacher-student relationships: Causes and consequences*. New York: Holt, Rhinehart & Wilson.
- Brown, R. (1965). *Social Psychology*. New York: Free Press.
- Brown, R. (1986). *Social psychology: The second edition*. New York: Free Press.
- Chaikin, A. L., Sigler, E., & Derlega, V. J. (1974). Nonverbal mediation of teacher expectancy effects. *Journal of Personality and Social Psychology*, 30, 144-149.
- Chaplin, W. F., Panter, A. T. (1993). Shared meaning and the convergence among observers' personality descriptions. *Journal of Personality*, 61, 553-586.
- Chelladurai, P. (1980). Leadership in sports organizations. *Canadian Journal of Applied Sport Sciences*, 5, 226-231.
- Chelladurai, P. (1984). Discrepancy between preferences and perceptions of leadership behavior and satisfaction of athletes in varying sports. *Journal of Sport Psychology*, 6, 27-41.
- Chelladurai, P. (1990). Leadership in sports: A review. *International Journal of Sport Psychology*, 21, 328-354.
- Chelladurai, P., & Carron, A. V. (1983). Athletic maturity and preferred leadership. *Journal of Sport Psychology*, 5, 371-380.

- Chelladurai, P., Malloy, D., Imamura, H., & Yamaguchi, Y. (1987). A cross-cultural study of preferred leadership in sports. *Canadian Journal of Applied Sport Sciences, 12*, 106-110.
- Christensen, D., & Rosenthal, R., (1982). Gender and nonverbal decoding skills as determinants of interpersonal expectancy effects. *Journal of Personality and Social Psychology, 42*, 75-87.
- Coladarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *Journal of Experimental Education, 60*, 323-337.
- Denham, C. H., & Michael, J. J. (1981). Teacher sense of efficacy: A definition and a model for further research. *Educational Research Quarterly, 5*, 39-63.
- Eckman, P., & Friesen, W. V. (1969). Nonverbal leakage and cues to deception. *Psychiatry, 32*, 88-106.
- Erdle, S., Murray, H. G., & Rushton, J. P. (1985). Personality, classroom behavior, and student ratings of college teaching effectiveness: A path analysis. *Journal of Educational Psychology, 77*, 394-407.
- Feltz, D. L., Chase, M. A., Moritz, S. E., & Sullivan, P. J. (1999). A conceptual model of coaching efficacy: Preliminary investigation and instrument development. *Journal of Educational Psychology, 91*, 765-776.
- Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology, 55*, 149-158.

- Funder, D. C., Kolar, D. C., & Blackman, M. C. (1995). Agreement among judges of personality: Interpersonal relations, similarity, and acquaintanceship. *Journal of Personality and Social Psychology, 69*, 656-672.
- Funder, D. C., & West, S. G. (1993). Consensus, self-other agreement, and accuracy in personality judgment: An introduction. *Journal of Personality, 61*, 457-467.
- Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76*, 569-582.
- Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday, Anchor Books.
- Goffman (1971). *Relations in Public*. New York: Basic Books.
- Good, T. L., & Brophy, I. E. (1987). *Looking in classrooms* (4th ed.). New York: Harper & Row.
- Gould (1987). Your role as a youth sports coach. In V. Seefeldt (Ed.), *Handbook for youth sport coaches* (pp. 17-32). Reston, VA: American Alliance for Health, Physical Education, Recreation, and Dance.
- Guskey, T. R., & Passoro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal, 31*, 627-643.
- Harris, M. J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: Thirty-one meta-analyses. *Psychological Bulletin, 97*, 363-386.
- Horn, T. S. (1992). Leadership effectiveness in the sport domain. In T. S. Horn (Ed.), *Advances in sport psychology* (pp. 181-199). Champaign, IL: Human Kinetics.
- Horn, T. (1984). Expectancy effects in the interscholastic athletic setting: Methodological considerations. *Journal of Sport Psychology, 6*, 60-76.

- Hoy, W. K., & Woolfolk, A. E. (1993). Teachers' sense of efficacy and the organizational health of schools. *Elementary School Journal, 93*, 355-372.
- John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality, 61*, 521-551.
- Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York, New York: Guilford Publications. Retrieved March 16, 2006 from <http://davidakenny.net/ip/assim.htm>
- Kenny, D. A. (2004). PERSON: A general model of interpersonal perception. *Personality and Social Psychology, 3*, 265-280.
- Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus and interpersonal perception: Acquaintanceship and the Big Five. *Psychological Bulletin, 116*, 245-258.
- Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L. (1992). Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology, 62*, 88-97.
- Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin, 102*, 390-402.
- Kenow, L. J., & Williams, J. M. (1992). Relationship between anxiety, self-confidence, and the evaluation of coaching behaviors. *The Sport Psychologist, 6*, 344-357.
- Kruglanski, A. W. (1989). The psychology of being "right": The problem of accuracy in social perception and cognition. *Psychological Bulletin, 106*, 395-409.

- Malloy, T. E., Agtaststein, F., Yaras, A., & Albright, L. (1997). Effects of communication, information overlap, and behavioral consistency on consensus in social perception, *73*, 270-280.
- Markland, R. D., & Martinek, T. J. (1988). Descriptive analysis of coach augmented feedback given to high school varsity female volleyball players. *Journal of Teaching in Physical Education*, *7*, 289-301.
- Martinek, T. J. (1988). Confirmation of a teacher expectancy model: Student perceptions and causal attributions of teaching behaviors. *Research Quarterly for Exercise and Sport*, *59*, 118-126.
- Martinek, T. J., & Johnson, S. (1979). Teacher expectations: Effects in dyadic interaction and self-concept in elementary physical education classes. *Research Quarterly for Exercise and Sport*, *50*, 60-70.
- Martinek, T. J., & Karper, W. (1984a). Multivariate relationships of specific impression cues with teacher expectations and dyadic interactions in elementary physical education classes. *Research Quarterly for Exercise and Sport*, *55*, 32-40.
- Martinek, T. J., & Karper, W. (1984b). The effects of competitive and noncompetitive instructional climates on teacher expectancy effects in elementary physical education classes. *Journal of Sport Psychology*, *6*, 408-421.
- Martinek, T. J., & Karper, W. (1986). Motor ability and instructional context: Effects on teacher expectation and dyadic interactions in elementary physical education classes. *Journal of Classroom Interaction*, *21*, 16-25.
- Maslach, C., & Jackson, S. E. (1981a). *The Maslach Burnout Inventory*. Research Edition. Palo Alto, CA: Consulting Psychologist Press.

- Maslach, C., & Jackson, S. E. (1981b). The measurement of experienced burnout. *Journal of Occupational Behavior*, 2, 99-113.
- Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality and Social Psychology*, 4, 681-691.
- Paunonen, S. V. (1989). Consensus in personality judgments: Moderating effects of target-rater acquaintanceship and behavior observability. *Journal of Personality and Social Psychology*, 56, 823-833.
- Price, M. S., & Weiss, M. R. (2000). Relationships among coach burnout, coach behaviors, and athletes' psychological responses. *The Sport Psychologist*, 14, 391-409.
- Rejeski, W., Darracott, C., & Hutskar, S. (1979). Pygmalion in youth sport: A field study. *Journal of Sport Psychology*, 1, 311-319.
- Riemer, H. A., & Chelladurai, P. (1995). Leadership and satisfaction in athletics. *Journal of Sport & Exercise Psychology*, 17, 276-293.
- Robins, R. W., & John, O. P. (1997). The quest for self-insight: Theory and research on accuracy and bias in self-perception. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 649-679). San Diego: CA: Academic Press.
- Rosenshine, B. (1971). *Teaching behaviors and student achievement*. London: National Foundation for Educational Research.
- Rosenshine, B., & Furst, N. (1973). The use of direct observations to study teaching. In R. Travers (Ed.), *Second handbook of research on teaching*. Chicago: Rand McNally.

- Rosenthal, R. (1966). *Experimenter effects in behavioral research*. Englewood Cliffs, NJ: Prentice-Hall.
- Rosenthal, R., & DePaulo, B. M. (1979). Sex differences in accommodation in nonverbal communication. In R. Rosenthal (Ed.), *Skill in nonverbal communication*. Cambridge, MA: Oelgeschlager, Gunn, & Hain.
- Rosenthal, R., & Jacobson, L. (1968). Pygmalion in the classroom: Teacher expectations and pupils intellectual development. New York: Holt, Rinhart & Winston.
- Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences*, 3, 377-415.
- Ryans, D. G. (1960). *Characteristics of teachers*. Washington DC: American Council on Education.
- Scherer, K. R. (1978). Personality inference from voice quality: The loud voice of extraversion. *European Journal of Social Psychology*, 8, 467-487.
- Sinclair, D. A., & Vealey, R. S. (1989). Effects of coaches' expectations and feedback on the self-perceptions of athletes. *Journal of Sport Behavior*, 12, 77-91.
- Smith, R. E., Smoll, F. L., & Curtis, B. (1979). Coach Effectiveness Training: A cognitive-behavioral approach to enhancing relationship skills in youth sport coaches. *Journal of Sport Psychology*, 1, 59-75.
- Smith, R. E., Smoll, F. L., & Hunt, E. (1977) A system for the behavioral assessment of athletic coaches. *Research Quarterly*, 48(2), 401-407.
- Snyder, M., Tanke, E., & Berscheid, E., (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, 44, 492-505.

- Solomon, G. B., Striegel, D. A., Eliot, J. F., Hean, S. N., Mass, J. L., & Wayda, V. K. (1996). The self-fulfilling prophecy in college basketball: Implications for effective coaching. *Journal of Applied Sport Psychology, 8*, 44-59.
- Spain, J.S., Eaton, L.G, & Funder, D.C. (2000). Perspectives on personality: The relative accuracy of self versus others for the prediction of emotion and behavior. *Journal of Personality, 68*, 837-867.
- Vealey, R. S., Armstrong, L., Comar, W., & Greenleaf, C. A. (1998). Influence of perceived coaching behaviors on burnout and competitive anxiety in female college athletes. *Journal of Applied Sport Psychology, 10*, 297-318.
- Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology, 57*, 120-128.
- Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*, Reading, MA: Addison-Wesley.
- Williams, J. M., Jerome, G. J., Kenow, L. J., Rogers, T., Sartain, T. A., & Darland, G. (2003). Factor structure of the coaching behavior questionnaire and its relationship to athlete variables. *Sport Psychologist, 17*, 16-35.
- Vincent, J. P., Friedman, L. C., Nugent, J., & Messerly, L, (1979). Demand characteristics in observations of marital interactions. *Journal of Consulting and Clinical Psychology, 47*, 557-566.
- Warr, P. B., & Knapper, C. (1968). *The perception of people and events*. London: John Wiley & Sons.

Weiss, M. R., & Friedrichs, W. D. (1986). The influence of leader behaviors, coach attributes, and institutional variables on performance and satisfaction of collegiate basketball teams. *Journal of Sport Psychology*, 8, 332-346.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology*, 10, 109-120.

APPENDIX C
INSTRUMENTATION



Student Coach Name _____

Your Name _____ **E-Mail** _____

Student Coaching Evaluation Tool

Please rate coach on next category as follows:

(1) Not at all (9) Very

Attributional Characteristics

Accepting	1	2	3	4	5	6	7	8	9
Active	1	2	3	4	5	6	7	8	9
Attentive	1	2	3	4	5	6	7	8	9
Competent	1	2	3	4	5	6	7	8	9
Confident	1	2	3	4	5	6	7	8	9
Dominant	1	2	3	4	5	6	7	8	9
Empathic	1	2	3	4	5	6	7	8	9
Enthusiastic	1	2	3	4	5	6	7	8	9
Honest	1	2	3	4	5	6	7	8	9
Likable	1	2	3	4	5	6	7	8	9
Optimistic	1	2	3	4	5	6	7	8	9
Professional	1	2	3	4	5	6	7	8	9
Supportive	1	2	3	4	5	6	7	8	9
Warm	1	2	3	4	5	6	7	8	9

Please rate coach on next category as follows:

(1) **Very Poor** (2) **Poor** (3) **Satisfactory** (4) **Good** (5) **Very Good**

Coaching/Coach

Overall, how would rate this coach?

1 2 3 4 5

Signature_____ Date_____

Thank you for your time!

If you have any questions, please feel free to contact me at lbryant@georgiasouthern.edu



Coach Self Evaluation Tool

Please rate self on next category as follows:

(1) Not at all (9) Very

Attributional Characteristics

Accepting	1	2	3	4	5	6	7	8	9
Active	1	2	3	4	5	6	7	8	9
Attentive	1	2	3	4	5	6	7	8	9
Competent	1	2	3	4	5	6	7	8	9
Confident	1	2	3	4	5	6	7	8	9
Dominant	1	2	3	4	5	6	7	8	9
Empathic	1	2	3	4	5	6	7	8	9
Enthusiastic	1	2	3	4	5	6	7	8	9
Honest	1	2	3	4	5	6	7	8	9
Likable	1	2	3	4	5	6	7	8	9
Optimistic	1	2	3	4	5	6	7	8	9
Professional	1	2	3	4	5	6	7	8	9
Supportive	1	2	3	4	5	6	7	8	9
Warm	1	2	3	4	5	6	7	8	9

Please rate self on next category as follows:

(1) Very Poor (2) Poor (3) Satisfactory (4) Good (5) Very Good

Coaching/Coach

Overall, how would you rate yourself as a coach?

1 2 3 4 5

COACH CLIP #1

Please rate coach on each attribute as follows:

(1) Not at all (9) Very

Attributional Characteristics

Accepting	1	2	3	4	5	6	7	8	9
Active	1	2	3	4	5	6	7	8	9
Attentive	1	2	3	4	5	6	7	8	9
Competent	1	2	3	4	5	6	7	8	9
Confident	1	2	3	4	5	6	7	8	9
Dominant	1	2	3	4	5	6	7	8	9
Empathic	1	2	3	4	5	6	7	8	9
Enthusiastic	1	2	3	4	5	6	7	8	9
Honest	1	2	3	4	5	6	7	8	9
Likable	1	2	3	4	5	6	7	8	9
Optimistic	1	2	3	4	5	6	7	8	9
Professional	1	2	3	4	5	6	7	8	9
Supportive	1	2	3	4	5	6	7	8	9
Warm	1	2	3	4	5	6	7	8	9

Please rate coach on next category as follows:

(1) Very Poor (2) Poor (3) Satisfactory (4) Good (5) Very Good

Coaching/Coach

Overall, how would rate this coach? 1 2 3 4 5

This document was created with Win2PDF available at <http://www.win2pdf.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.
This page will not be added after purchasing Win2PDF.