

Research Paper: Sample Critique

Jamber, E. A., & Zhang, J. J. (1997). Investigating leadership, gender, and coaching level using the Revised Leadership for Sport Scale. *Journal of Sport Behavior*, 20, 313-322.

The purpose of the study was to determine possible differences in leadership behaviors, using the Revised Leadership for Sport Scale (RLSS), between male and female coaches and among different coaching levels. The researchers submitted two hypotheses. The first hypothesis was that male and female coaches would respond differently to the RLSS in overall leadership behaviors. The second hypothesis was that differences on the RLSS would occur among coaching levels: junior high, high school, and college.

The sample was nonrandom, including 162 coaches that were chosen on a volunteer basis. Within the sample, 118 (0.73) of the coaches were male, while 44 (0.27) were female. With regard to coaching level, 25 (0.15) were junior high coaches, 99 (0.61) high school, and 38 (0.24) at the college level. While this is a good sample size, the problem lies with the distribution of the sample. The sample number for junior high coaches, in particular, is rather low. A larger sample with regard to all categories would have aided in the data analysis, particularly when looking for possible interactions between gender and coaching level.

The instrument utilized was the Revised Leadership for Sport Scale (RLSS) developed by Zhang, Jensen, and Mann in 1996. This scale is used to measure six leadership behaviors: training and instruction, democratic, autocratic, social support, positive feedback, and situational consideration. The scale uses 60 statements, which were preceded by “In coaching, I:” A Likert scale was then given for each statement: 1 = never; 2 = seldom; 3 = occasionally; 4 = often; and 5 = always. This produced an ordinal level data set. Scales were administered in a number of environmental settings: classrooms, gymnasiums, practice fields, and offices. The internal

consistency for each section was calculated: 0.84 for training and instruction; 0.66 for democratic; 0.70 for autocratic; 0.52 for social support; 0.78 for positive feedback; and 0.69 for situational consideration. There was no information, however, regarding the validity of the RLSS.

A MANOVA was used to analyze the data for differences between male and female coaches with regard to leadership behaviors. This is not consistent with the type of data collected. The RLSS used a Likert scale (ordinal), yet a MANOVA would be most applicable for normally distributed, quantitative data. The analysis showed there were no significant differences between male and female coaches in overall leadership behaviors. When the six leadership styles were examined separately, there was a significant difference in social support between males and females. In general, females scored much higher than did the male coaches.

A MANOVA was also used to examine the data for differences between the three levels of coaching (junior high, high school, and college) with regard to leadership behavior in general. There were significant differences between the three levels. When breaking down the six behaviors and examining them individually, an ANOVA was used to analyze the data. Again, because the data for the RLSS is ordinal, an ANOVA is not the best analysis tool. The three coaching levels scored differently on three of the six behaviors: democratic behaviors, training and instruction, and social support. High school coaches scored much higher than college level coaches in democratic behavior. Junior high coaches were significantly lower in training and instruction than either high school or college coaches. Junior high coaches also demonstrated a lesser degree of social support than either the high school or college coaches.

A MANOVA was again used to analyze the data for any interaction between gender and coaching level with regard to overall leadership behavior. Once again, a better analysis method could have been chosen based on the nature of the data collected. The results indicated no significant interactions.

The ecological generalizability for the study is fairly high. The surveys were mailed out, and returned on a volunteer basis. However, due to the nonrandom nature of the sample, the results would not generalizable beyond the 162 participants in the study. There was no effect size is listed for the study.

In order to reduce threats to internal validity, the participants were asked to respond honestly and confidentiality was stressed so that the “coaches might feel more at ease in responding.” No other efforts were indicated.

The researchers mention that the scales were given in a variety of settings. This could present a threat to the internal validity in that participants might not have been entirely focused on completing the scale, but instead on coordinating practice, completing paperwork, etc. There are a number of other factors that could affect the internal validity of the study, yet were not addressed by the researchers. Coaching experience would greatly affect the responses of the participants, yet this was not considered in the study. The gender of the athletes may be a contributing factor to the coaches’ responses. It is not unreasonable to suppose that coaches of female athletes, particularly at the junior high and high school levels, will demonstrate more social support than those of male athletes. The nature of the sport could also be critical. Certain coaching styles are more applicable for individual sports (wrestling, track, and tennis) than for team sports (football, soccer, and basketball). The socioeconomics and population of the school itself could play a factor. Certain schools have better athletes and programs in a particular sport, while others may not be able to field a winning team. In addition, at the high school level, coaches are occasionally asked/forced to work with a program they have no knowledge of or desire to coach due to staffing shortages. This could dramatically influence a coach’s response to the scale questions. The history of the program as well as the individual coach’s personal coaching history could greatly influence responses. If the program has had several losing seasons in a row,

perhaps the attitude of the coach could be different than that of a coach who has recently won a state title.

An additional set of questions regarding the personal history of the coach in question could have helped reduce many of these threats. With additional information, the researchers may have been able to use a modified matching system when analyzing the results. By increasing the number of independent variables to include things such as coaching experience and gender of the athletes, the researchers could have reduced some of the potential threats to internal validity. In addition, bringing coaches together to a common setting could have reduced location threat. Coaches meet seasonally for clinics. Perhaps obtaining permission to administer the survey during these meetings would have been possible. It would have also been possible to actually go to individual schools and meet with the coaches as a group to administer surveys. This method would have given a good cross-section of gender and coaching experience for a variety of sports.

While the study has merit, **the methods need to be re-evaluated.** The power of the study needs to be increased by obtaining a larger sample size. The numerous potential threats to internal validity need to be addressed and minimized where possible. It would also be helpful to be given data regarding the validity of the RLSS. **Without these, it is impossible to evaluate the potential meaningfulness of this study.**