

## Use of Secondary Data in Sport Management Research

Ming Li, Ohio University

Heather Lawrence, Ohio University

### Session 5: Research Presentation (25-minute)

Thursday, May 31, 2007

11:30 AM - 11:55 AM

The review of the articles published in the *Journal of Sport Management* since 1992 and in *Sport Marketing Quarterly* since 2003 indicated that most of the research conducted in sport management involves the collection of new data, or the primary data to solve a particular problem. Nevertheless, sport management researchers sometimes may turn their attention to secondary data for help to solve their research problems. Secondary data refer to the type of data that have previously been gathered and that might be relevant to the problem at hand (McDaniel & Gates, 1991). The purpose of this study is two-fold: (a) to identify the common ways that sport management researchers have used the secondary data design in solving their research problems and (b) to demonstrate the use of a secondary data set in sport management research with a case. Approximately 250 articles were reviewed, 25 of which or 10% used the secondary data design.

The results of this review indicate that sport management researchers have used secondary data design primarily in three different ways: to answer a research question, to clarify and understand a situation, and to build a statistical model. Researchers are often use secondary data to either answer a research question or approve a hypothesis (e.g., Ahlstrom, Si, & Kennelly, 1999; Baird, 2004; Fortunato, 2004; McDonald & Rascher, 2000; O'Brien & Slack, 2004; Orders & Chelladurai, 1994; Stinson and Howard, 2004; Sack, Singh, & Triel, 2005; Zygmunt & Leadley, 2005). "The first tenet of data gathering among researchers is to exhaust all sources of secondary data before engaging in a search for primary data" (Wrenn et al., 2002). In fact, many research questions can simply be answered with the analysis of the data collected from a secondary source. For example, to understand whether or not winning has a significant impact on alumni giving, Stinson and Howard (2004) examine the secondary data on donors making gifts of \$1,000 or more between 1994 and 2002 to the Annual Giving Program at the University of Oregon. They found "in almost every year, alumni giving to athletics has increased with an associated increase in success by high profile intercollegiate athletic teams..." (p. 136). In a study conducted by McDonald & Rascher (2000), the secondary data design was used to answer a research question, "Do promotions work" in increasing fan attendance. The second major use of secondary data by sport management researchers is to clarify and understand a situation or fact finding (e.g., Crompton, Howard, & Var, 2003; Howard, 1992; Taylor, 2003). For instance, to examine the issues of cultural diversity in women's sport, Taylor (2003) first used secondary data to generate exploratory information about the participation rates in sport and physical activities among women from minority ethnic backgrounds and Anglo-Australian women to verify that the former group of women had a lower participation rate. To further confirm the findings from the secondary data research, the researcher conducted a survey to investigate the issue with a specific sport-netball. The empirical data confirmed that the levels of female participation from culturally diverse backgrounds in netball were still significantly below those of Anglo-Australian females.

Secondary data are used most frequently (12 out of the 20 identified articles, or 60%) by sport management researchers to build a statistical model that has the ability (a) to specify and make prediction about relationships between variables, or (b) to approve a conceptual model (e.g., Agthe & Billings, 2000; Boronico & Newbert, 1999; DeSchraver & Stotlar, 1996; Dobson & Gerrard, 1999; Fizek & D'itri, 1996; Fortunato, 2004; Gerrard, 2005; Hadley & Gustafson, 1991; Holbrook & Shultz, 1996; Maxcy & Mondello, 2006; Rascher & Rascher, 2004; Todd, Crook, & Barilla, 2005; Yokum, Gonzalez, & Badgett, 2006). For example, Hadley & Gustafson (1991) used secondary data to establish a single-equation model of baseball salaries. The study found that such variables as years of major league service, eligibility for arbitration, etc. had a great impact on players' salaries in major league baseball.

To demonstrate how to use secondary data in sport management research, two secondary data sets on intercollegiate athletics were obtained. The first one contains intercollegiate financial data and was extracted from the Equity in Athletics Disclosure Act (EADA) report. The other source was the Directors Cup finishes established by the National Association of Collegiate Director of Athletics (NACDA). The Directors Cup points were based national finishes. Two research questions were then raised: Does a relationship exist between financial variables and Directors Cup points within the same academic year? What are the predictors of the Directors Cup points? A correlation analysis was performed. The results indicated positive correlations between Directors Cup points and overall revenues ( $r=.497, p<.01$ ), expenses ( $r=.502, p<.01$ ), expense per female athlete ( $r=.411, p<.01$ ), expense per male athlete ( $r=.507, p<.01$ ), per athlete operating expenses by gender ( $r=.378, p<.01$ ), women's recruiting expenses ( $r=.375, p<.01$ ), and men's recruiting expenses ( $r=.438, p<.01$ ). The only variable that was not significantly correlated with the Directors Cup points was the difference between revenues and expenses. To answer the question of what are the predictors of Directors Cup points, a stepwise multiple regression analysis was performed with the Directors Cup points as the dependent

## 2007 North American Society for Sport Management Conference (NASSM 2007)

variable, and financial variables along with the number of male and female athletes as independent variables. Two predictors, the expenses per male athlete (unallocated \$ excluded) and the number of male athletes were identified in a regression model with a  $R^2 = .369$  and  $F$  value = 20.783 was significant at .01 level.

The researcher that uses the secondary data design should evaluate the data before attempting to use them. The evaluation involves a number of steps. The first step is to verify the data: where the data came from, who was the individual or organization responsible for collecting the data, and was the individual or organization reputable and creditable. "The source of the secondary data is the key to its accuracy" (McDaniel & Gates, 1991, p. 124). The EADA data were generated with a self-report mechanism, which may have created inconsistency in categorizing financial information. For example, the athletic department in one institution may be responsible for utility bills as an athletic expenditure while utility expenses may fall under the general operations of another institution as opposed to an athletic expense. The next step is to evaluate the methodology used in data collection. In this case, once institutions filed the report online, there is not a mechanism to ensure the submitted data were accurate.

The integration of secondary data into sport management research is a research design that has been utilized by sport management researchers. Nevertheless, while enjoying the convenience in accessing secondary data, researchers should approach them with care.