Paper/Pencil Versus Online Data Collection:
An Exploratory Study

Peter Ward
Taralyn Clark
Ramon Zabriskie
Brigham Young University

Trevor Morris
Utah Valley University

Abstract

Usage of online data collection methods is increasing in leisure research. Some potential benefits to using online methods over traditional paper/pencil techniques include financial savings and easier access to large populations. Disadvantages, however, include difficulty in sample selection and variations of the instruments' reliability. This study explores how subject responses potentially differ when collecting data online versus paper/pencil for six instruments commonly used within leisure research. A repeated measure design with paired sample t-tests and HLM was used with 207 college students to compare these methods of data collection. Responses differed between methods on three of the tested instruments. A general pattern was found suggesting participants perceived their anonymity was better protected when completing online questionnaires.

KEYWORDS: Data collection, online, anonymity, methods

Peter Ward is an assistant professor, Taralyn Clark received her MS, and Ramon Zabriskie is a professor in the Department of Recreation Management at Brigham Young University.
Trevor Morris is and instructor in the Department of Psychology at Utah Valley University.
Address correspondence to Peter Ward, Recreation Management and Youth Leadership Department, Brigham Young University, W427 Tanner Building, Provo, Utah, 84602, peter_ward@byu.edu.

Online research methods have been used with leisure studies over 30 times in top-tier leisure journals from January 2000 to present (Dodd, Zabriskie, Widmer, & Eggett, 2009; Gladwell, Dorwart, Stone, & Hammond, 2010; Li & Petrick, 2010; Swinton, Freeman, Zabriskie, & Fields, 2008; Tu, Chen, Wang, & Lin, 2007). Although the trend of implementing this technique is growing in many disciplines, particularly strong growth has been evident in social and leisure sciences (Cronk & West, 2002; Granello & Wheaton, 2004; Hardre, Crowson, Xie, & Ly, 2007; Rademacher & Lippke, 2007; Wright & Schwager, 2008). For example, in 2001 only one article using online methods was published in a major leisure journal, whereas three major leisure journals published 10 studies using online methods in 2009 and the beginning of 2010. Increased usage of online methods is apparent in leisure research, yet the investigation into whether online methods and traditional methods produce the same participant responses has not been explored in a leisure setting.

Benefits of online data collection lead some researchers to posit these methods will continue to grow and may even replace traditional paper data collection (Lefever, Dal, & Matthiasdottir, 2007). Among reported advantages of online data collection are financial savings, fewer time limitations, more accurate data collection, easier access to large populations, and increased anonymity for study participants (Aluja, Rossier, & Zuckerman, 2007; Buchanan, 2002; Cronk & West, 2002; Davis, 1999; Miller et al., 2002; Riva, Teruzzi, & Anolli, 2003). In spite of these advantages, several disadvantages to online data collection have also been noted, including difficulty in sampling select participants and variation of the instruments' reliability when compared to traditional data collection methods (Granello & Wheaton, 2004; Lefever et al., 2007; Schillewaert & Meulemeester, 2005; Topp & Pawloski, 2002).

As utilization of online research methods increases, studies continuing to examine this trend are needed (Lonsdale, Hodge, & Rose, 2006; Raat, Mangunkusumo, Landgraf, Kloek, & Brug, 2007), specifically studies detailing the performance of leisure instruments as to how they compare to their paper/pencil counterparts. Some have advocated testing an instrument online is essential regardless of previous paper/pencil results (Aluja et al., 2007; Buchanan, 2002; Buchanan et al., 2005; Davis, 1999; Hewson & Charlton, 2005; Touvier et al., 2010). Therefore, the purpose of this study was to compare the use of online data collection with the paper/pencil versions of six instruments often used in leisure research.

Review of Literature

Growth in the use of online data collection procedures is a direct result of many people having easy access to the World Wide Web (Granello & Wheaton, 2004). However, social science researchers are sometimes using traditional paper/pencil instruments adapted for online use without understanding the implications they may have for the study's results. Akin to other social sciences, previous leisure studies utilizing online versions of traditional instruments cannot be generalized to all instruments (Buchanan, 2002). Therefore the literature review will discuss what is unique about leisure research, advantages and disadvantages of online data collection, and further related considerations.

Leisure Research

Leisure instruments need to be studied independently of other social science instruments. Participating in leisure is different than nearly any other activity people do or mind-set they are in (Kelly, 1996; Mannell & Kleiber, 1997). Leisure invites the use of personal freedom and is an expression of how individuals intrinsically choose to exercise that freedom (Nash, 1953). In light of this, Mannell and Kleiber (1997) suggest people's leisure behavior may be different than their behavior in situations encountered more regularly. Research also indicates leisure behavior may impact quality of life more powerfully than any other behavior (Kelly, 1996; Mactavish & Schleien, 1998). The outcome participants seek from leisure is a self-chosen experience with meaningful and significant implications.

One example of the impact leisure has on an individual's quality of life is illustrated through examining leisure's influence on personal and family relationships. A Canadian Parks and Recreation Association study reported, "In modern society, leisure is the single most important force developing cohesive, healthy relationships between husbands and wives and between parents and their children" (as cited in Hornig, 2005, p. 48). Additionally, research consistently shows leisure enhances one's ability to cope, improves life satisfaction, and contributes to greater family cohesiveness and stability, thereby strengthening families and improving family functioning (Bocarro & Sable, 2003; Greeff & Leroux, 1999; Hornberger, Zabriskie, & Freeman, 2010; Kimball & Freysinger, 2003; Mactavish & Schleien, 1998; Orthner, 1998; Orthner & Mancini, 1990; Palmer, Freeman, & Zabriskie, 2007; Zabriskie, 2001). These inherent aspects of leisure cause it to be fundamentally different than other social sciences.

Accurately understanding how people spend and evaluate their leisure time, and even grow through leisure, is a unique subject that cannot be folded in with other social sciences. Although leisure is informed by other social sciences, it is important to study leisure as its own discipline. Leisure scholars can use social sciences as a guide for research methods and data collection procedures, but they need to independently test to see if these methods are valid for data collection within the leisure arena. Therefore, in light of increasing trends to conduct leisure research online, it is important to explore possible differences between the use of leisure instruments administered online versus traditional methods to ensure leisure constructs are being measured accurately and results can be generalized and applied.

Advantages and Disadvantages

In the United States, nearly 220 million individuals, or 73% of the total population, have Internet access (Internet World Stats, 2009). While the trend of online data collection is growing in leisure research, this technique still yields an opportunity that is being underutilized by scholars. The capacity to reach a broad span of diverse individuals is made possible due to modern online data collection methods. Before using this approach, it is important to understand the distinct advantages and disadvantages online data collection techniques offer.

Advantages. Many researchers suggest the Internet can be and is used as a viable research tool. Some benefits this tool may provide include lower research

costs, less time spent collecting data, more clean and accurate data collection, access to a large and diverse population, and enhanced experience and anonymity for participants (Ahern, 2005; Aluja et al., 2007; Ballard & Prine, 2002; Buchanan, 2002; Buchanan et al., 2005; Cronk & West, 2002; Davis, 1999; Hewson & Charlton, 2005; Lonsdale et al., 2006; Lutner et al., 1991; Miller et al., 2002; Pettit, 2002; Raat et al., 2007; Riva et al., 2003). Lower data collection costs are one of the main advantages to online data collection. McDonald and Adam (2003) found traditional mail response research was twice as expensive compared to online surveys. Others report a financial savings between 20% and 80% when compared to other data collection techniques such as paper survey administration (Granello & Wheaton, 2004; Illieva, Baron, & Healey, 2002; Lefever et al., 2007; Rhodes, Bowie, & Hergenrather, 2003). These savings are valuable because they allow researchers to allocate resources to other projects or areas of interest, promoting a broadening of research. Paper/pencil questionnaires require paper costs, printing costs, and data entry costs. If surveys are administered in a classroom or similar setting, costs could possibly entail paying and training administrators and travel, as well as costs associated with postal mailings.

Along with saving money, scholars can save time by collecting data online. McDonald and Adam (2003) found a quicker response time with online data collection compared to postal data collection. More specifically, studies show response time was reduced to as little as two or three days (Granello & Wheaton, 2004) versus weeks and often months due to the streamlined electronic return process over slower traditional data collection (Lefever et al., 2007). In addition to actual data collection time savings, most online surveys can be programmed to input data directly into statistical packages. This saves the researcher time as well as eliminates error associated with data entry. Moreover, the ability to access large amounts of participants within a short time frame is a major advantage to using online data collection (Granello & Wheaton, 2004; Illieva et al., 2002; Lefever et al., 2007; McDonald & Adam, 2003). The Internet is available in much of the world and can be utilized to gather data from foreign countries without travel expense (Touvier et al., 2010).

Participants' experience can also be enhanced by technology used to administer surveys. Ellis and Rossman (2008) posit value can be increased by intentionally enhancing participant experience in the leisure industry. This same logic can be applied to participants' experience when completing surveys. Intentionally designing questionnaires can increase participants' involvement with the survey and thus create an experience for them during the data collection process. The ability to design this type of interactive survey is greatly augmented with the use of electronic media. Different color, sound, graphics, and video have been used to provide variety as well as to enhance or clarify (Dillman, 2000; Granello & Wheaton, 2004; Topp & Pawloski, 2002). Online methods enable participants to freely take the survey at their convenience, allowing participants the autonomy to schedule the data collection experience for a time and location that is best for them (Rhodes et al., 2003). Participant experience can also be enhanced when using online collection methods as setting and administration methods are held constant (Rhodes et al., 2003). Touvier et al. (2010) report study participants preferred completing

instruments online to paper versions. Using online techniques can enhance participant experience and also improve participants' perception of anonymity, possibly encouraging honesty in responses (Ahern, 2005).

Perception of anonymity when a person is online is exhibited in studies regarding Internet addiction (Cooper, Delmonico, & Burg, 2000; Demmel, 2002; Wong, 2010), social anxiety (Shepherd & Edelmann, 2005), and online counseling (Young, 2005). Individuals with Internet addictions such as online sex and intimate relationships, gambling, and web cruising report engaging in online activities they would not perform in a personal face-to-face social interaction. Additionally, accessibility and affordability of online activities encourage nefarious behavior (Cooper et al., 2000; Wong, 2010). Seemingly on the opposite side of the spectrum, individuals also report positive aspects of Internet anonymity, including comfort for those suffering from social anxiety (Shepherd & Edelmann, 2005). In a study of client attitudes toward online counseling, 96% of participants reported seeking online counseling due to anonymity, with 71% adding convenience to a list of reasons (Young, 2005). Well-established norms of online anonymity likely impact online research as study participants may feel safer in reporting sensitive issues, such as involvement in socially undesirable thoughts and behavior (Wang et al., 2005). Such findings illustrate how participant perception of anonymity is altered due to method of data collection. This fluctuation may impact study results, especially if sensitive questions are being asked.

Many researchers claim there are no differences between participants' responses in online survey completion and paper/pencil surveys, while other researchers claim participants respond more openly and honestly to online surveys (Lefever et al., 2007; Rhodes et al., 2003). One study compared phone surveys to online surveys and noted respondents had a lower mean score on positively worded items to the online surveys, suggesting respondents may offer more direct answers to questions in an online format (Roster, Rogers, Albaum, & Klein, 2004). Furthermore, adolescents are more likely to provide truthful responses to sensitive questions asked online than in other survey formats (Rhodes et al., 2003; Wright & Schwager, 2008). This may be because adolescents are striving to give the socially desired response when a person is administering the survey (Klein, Havens, & Thomas, 2009) or because adolescents mistrust a physical person who can trace the survey back to them (Wright, Aquilino, & Supple, 1998). This is demonstrated in Wright et al.'s (1998) study when researchers asked adolescents about number of sexual partners, unprotected sexual practice, and illegal drug use. Participants were told all responses were confidential and no one would know who offered what responses. Adolescents, however, reported much higher participation on the mentioned criteria on the online survey when compared to the paper and pencil survey. Adolescents may not be offering a completely truthful response if they believe their answers could be linked to them because they do not want to be judged or held accountable for their actions. This phenomenon possibly accounts for some differences in adolescents' responses to online compared to paper surveys. Additionally, social desirability forces may be perceived by respondents as lower in an online survey, resulting in more truthful answers (Johnson, 1999). Because "it is possible that perceived anonymity is more important than real anonymity"

(Buchanan, 2002, p. 150), perceived anonymity might be an important factor in determining reliability of online instruments, thereby meriting further study.

Disadvantages. While technology and Internet use may have great advantages, they can also pose disadvantages. For example, distributing online surveys by sending a link to a website's questionnaire through e-mail invitation is popular and very convenient, but may have limitations (Duffy, 2002). Some participants may have multiple e-mail addresses they no longer use. Others may use e-mail filters to sort junk and unwanted mail from what is important and wanted. These issues may limit the sample a researcher is trying to reach, therefore biasing the sample and lowering response rates. Additionally, the desired sample must be taken under consideration when determining data collection methods as problems may arise due to age-related barriers or other accessibility concerns (Klovning, Sandvik, & Hunskaar, 2009).

Technology presents challenges for participants and researchers alike, though these challenges are constantly changing. Topp and Pawloski (2002) address the issues of creating and maintaining passwords, access to listservs, and maintenance of server speeds associated with gathering data online, which can all be expensive. Acquiring the physical technology and know-how necessary for online data collection can also be too complex for the researcher wanting to create an online survey (McDonald & Adam, 2003), though survey programs such as Qualtrics and Survey Monkey have greatly reduced this disadvantage by increasing ease of creating and distributing online surveys.

Additionally, the researcher must address several key issues including generalizability, ethical issues, and reliability of the instrument. Granello and Wheaton (2004) argued Internet users are generally white, male, married, and educated, although the typical person to complete online surveys does not necessarily match this description. In a study completed with 16- to 19-year-olds, 65% of volunteer respondents were female, suggesting a need to pay attention to respondent characteristics in deciding whether to use online data collection (Lefever et al., 2007). The fit between target population and sample is not unique to online surveys, but still must be addressed in order to increase generalizability of online surveys.

Ethical concerns also need to be addressed with online research (Ahern, 2005). Ethical practices of both the researcher and the respondent have been brought into review (Rhodes et al., 2003). One key issue is participant anonymity. Participants believe their anonymity is protected while responding to online surveys, but it may not be. This method provides avenues in which researchers could potentially identify participants and possibly sell their information or use it for other gain. In addition, participants may use electronic means to quickly complete the same survey again and again to receive an incentive multiple times, leading to erroneous results. Obtaining parental consent is also a concern, as methods to accomplish this are not well established and often cumbersome. Some current practices include providing a consent form that must be signed and returned via postal mail or fax, requiring a parent to use a credit card with an online transaction, having parents call a phone number to obtain an authorization code, and using e-mail with an associated pin number that authorizes access to the survey. A

critical ethical consideration researchers should consider is how easy it may be to provide services to respondents who may have been harmed from participating in the study. Because the participant is far removed from the researcher, it may often be extremely difficult to provide restorative services.

An often-mentioned disadvantage associated with online data collection, and a main concern in this study, is the psychometric properties for online instruments adapted from paper and pencil versions (Granello & Wheaton, 2004; Schillewaert & Meulemeester, 2005; Topp & Pawloski, 2002). Granello and Wheaton (2004) argue little is known about the psychometric implications of transferring a survey from paper/pencil to an electronic format. Previous studies suggest online instruments need to be tested for reliability regardless of paper/pencil results that may have been previously received (Aluja et al., 2007; Buchanan, 2002; Buchanan et al., 2005; Davis, 1999; Hewson & Charlton, 2005). "Although there is evidence that online and traditional versions of the same test can measure the same constructs, there is also evidence that the instruments are not always identical" (Buchanan, 2002, p. 150). Instruments can be the same in appearance and structure but differ in the way participants perceive them. For example, participants may think an instrument offers more anonymity or convenience solely based on the method of data collection, thereby altering their responses.

Researchers, especially in psychology and health-related fields, have responded to the need for psychometric testing by conducting studies comparing online and paper/pencil versions of various instruments measuring personality traits, perceptions on community policing, memory, athlete burnout, alcohol measures, and health-related quality of life (Aluja et al., 2007; Ballard & Prine, 2002; Buchanan et al., 2005; Cronk & West, 2002; Davis, 1999; Hewson & Charlton, 2005; Lonsdale et al., 2006; Miller et al., 2002; Pettit, 2002; Raat et al., 2007; & Riva et al., 2003). Such studies need to be conducted with each individual instrument, as identical psychometric characteristics cannot be assumed (Buchanan, 2002). Although most studies report psychometric statistics in online instruments, Buchanan et al. (2005) found these psychometric statistics were not the same for the online and the paper and pencil version for the multidimensional instruments studied, and they called for more related research. For example, Woolhouse and Myers (1999) compared paper and pencil with online versions of a multidimensional personality inventory and detected differences existed in the factor structure of inventories. Some items loaded on different factors when the two means of data collection were compared. Bachana, Goldberg, and Johnson (1999) found similar results using a different multidimensional personality inventory. \ Thus, the online version of these tests did not appear to have the same psychometric properties as the paper and pencil instrument on which they were based. \ Ostensibly, further research is needed to reach a definitive answer with regard to the difference between electronic sampling techniques versus paper/pencil methods.

Further Considerations

The established trend in social behavior sciences is shifting toward using technology and online methods to conduct survey research. The same trend is emerging among leisure researchers. As leisure researchers consider collecting online data, it is important for them to understand the advantages, the disadvantages,

and the nature of responses they may receive from each method. Although use of online data collection methods is a growing trend in leisure research (e.g., Agate, Zabriskie, Agate, & Poff, 2009; Hornberger et al., 2010), no studies establishing the reliability of commonly used leisure survey instruments have previously been conducted. Previous studies argue for the need to test all instruments being adapted from a paper/pencil to an online format (Buchanan et al., 2005; Lonsdale et al., 2006).

Researchers also suggest a need for comparisons between online and paper/pencil instruments to be taken by the same participants using a repeated measure design, as this has not been previously done (Lonsdale et al., 2006; Raat et al., 2007). Typically a convenience sample of participants takes the online version, while others take the paper/pencil version only. Thus, a true comparison of the nested effect between individuals has not been investigated. Similarly, Lutner et al. (1991) used a repeated measures design in comparing automated with traditional interviews, suggesting more repeated measures designs are conducted. As such, the purpose of this study was to implement a repeated measures design to compare the use of online data collection with the paper/pencil versions of six instruments often used in leisure research.

Methods

Sample

Data were collected from a convenience sample of university students on two different campuses in the western United States. Students within medium-sized introductory classes of psychology and leisure were asked to participate via a brief in-class introduction by a member of the research team. Two different academic topics were chosen in order to capture a more diverse sample. One psychology class met the requirement for a general education class and attracted students from many different academic interests. An introduction to leisure class consisted of students who were leisure majors, nonprofit management minors, or students exploring the leisure and nonprofit academic programs. No participants were enrolled in both classes.

The sample consisted of 207 participants. After the researchers cleaned and screened the data, 141 participants completed both the online and paper/pencil questionnaires, while the remainder of participants (*n* = 66) only completed either the online or the paper/pencil questionnaire. Of the 207 respondents, 182 completed the demographic section. Therefore, demographic descriptive statistics represent 87.9% of the total sample. Additionally, eight participants did not complete all sections of the online questionnaire, resulting in variation of sample size between the instruments. Participants were dropped from the sample due to grossly incomplete questionnaires or for providing responses that were obviously not valid (i.e., marking the same response for all items).

The sample consisted of 46 males and 138 females; one of the subjects did not provide a gender. The average year in college was a second semester sophomore, with 46 freshman, 59 sophomores, 41 juniors, and 39 seniors. Participants' average personal income was between $5,000 and $10,000 annually, with average

family income between $60,000 and $70,000 per year. Nearly 80% of the students were Caucasian, followed by almost 4% Hispanic. Remaining participants varied across African American, Asian, and others. Just over 73% of the sample was single and had never married with 13% married and the rest either divorced or other. Participants reported spending an average of 10–12 hours per week on the computer. All participants said they spent at least 1–2 hours a week on the computer with the most time being greater than 31 hours a week. On a scale of 0 (*no computer skills*) to 8 (*expert computer skills*) the sample reported an average computer skill level of 4.96 (*SD* = 1.23).

Data Collection Procedures

Data were collected via two formats in two phases. First, a traditional paper/pencil instrument was developed that students were asked to take in their respective classroom settings. The second format was an online version via Qualtrics. A link to the online version of the questionnaire was e-mailed to students via the respective universities' electronic classroom management system. Students then accessed the questionnaire by clicking on the link and completing the questionnaire while online.

These two methods of data collection were selected because paper/pencil instruments surveying university students are typically administered in controlled settings, often in the classroom with a researcher present (Alexandris, Funk, & Pritchard, 2011; Breunig, O'Connell, Todd, Anderson, & Young, 2010; Gallant, Smale, & Arai, 2010; Wang & Walker, 2011), while online instruments in the leisure field are typically administered in uncontrolled settings over a span of time (Agate et al., 2009; Bruton et al., 2011; Dodd et al., 2009; Gladwell et al., 2010; Jun & Kyle, 2011; Mitas, Qian, Yarnal, & Kerstetter, 2011; Mowen, Payne, Orsega-Smith, & Godbey, 2009; Mulvaney, 2011; Poff, Zabriskie, & Townsend, 2010; Salk & Schneider, 2009). Furthermore, Woolhouse and Myers (1999) state collecting data online and comparing it to other data collection methods needs to be done in a realistic environment where participants would actually take the survey. Hardre et al. (2007) suggest paper/pencil versus online and computer-based future research be taken out of the laboratory to more realistic environments, accounting for how respondents will actually participate in the research. Furthermore, they elaborate that allowances need to be made for not being able to control the research setting. Potential challenges such as extraneous (e.g., distraction, environmental cues, technical variability between different hardware and software) and temporary (e.g., fatigue, altered states of mind) factors that may influence responses must be left in the equation when comparing methods. In this study, researchers modeled the actual participant experience as closely as possible in order to make the two data collection methods realistic to how it would be done in real data collection situations (e.g., paper and pencil sit down at a table or desk and complete with the researcher present compared to online done at home at participant's convenience). By collecting data consistent with how it is traditionally done, it was believed a more parallel comparison could be made that would have validity to the leisure researcher.

Great effort was made to reduce overall differences between the online and paper and pencil version of the questionnaire. Questionnaire format was the same

between the two instruments. Both instruments were very similar to their counterparts with the same use of spacing, font size and type, item ordering, overall structure, and color. Every effort was made to keep the instruments as identical as possible so the only difference between them was participants taking them online or via paper and pencil. Participants were asked not to leave any individual items blank in the paper and pencil and online versions.

In phase one, roughly half of the subjects received the online format first (*n* = 103), while remaining participants received the paper/pencil format. Efforts were made to randomize who received which format first, but in order to manage the study, randomization was done on the classroom level. Thus, all participants within a single class received either the online version or paper/pencil version first. Approximately five weeks later students were asked to complete the other format. It was believed 5 weeks was sufficient time for students to forget their responses to questions, thus reducing test–retest bias. This assumption was made based on Raat et al.'s (2007) study that only allowed for 2 weeks between participants taking the questionnaire as sufficient time to forget previous answers. The same procedures utilized to administer questionnaires in the first phase were used when administering the questionnaire in the second phase. Students for both the first and second phase of data collection were given 10 days to complete the online survey. E-mail reminders were sent to participants twice during this time. After 10 days, no more online questionnaires were accepted. Participants spent approximately 20 to 30 minutes completing the paper/pencil version. No data were collected on how long participants took to complete the online version, but in pilot testing participants were observed to take about the same amount of time.

Instrumentation

Six instruments relating to leisure and social psychology were utilized to explore the differences between data collected via online questionnaires and data collected via paper questionnaires. Instruments included Marlowe-Crowne (Crowne & Marlowe, 1960), Morally Debatable Behaviors (Harding, Phillips, & Fogarty, 1986), The Way I Feel About Myself (Piers, 1984), Leisure Satisfaction Measure (Beard & Ragheb, 1980), Leisure Boredom (Iso-Ahola & Weissinger, 1987), and Satisfaction With Life Scale (Diener, Emmons, Larsen, & Griffin, 1985). These questionnaires were chosen because they are firmly established paper/pencil instruments commonly used in leisure and social science research.

The instruments were placed in parallel order across both methods of data collection. The Marlowe-Crowne scale was followed by Morally Debatable Behaviors scale, How I Feel About Myself, Leisure Satisfaction Measure, Leisure Boredom scale, and Satisfaction With Life Scale and was then followed by demographic questions. A parallel order of instrument completion was desired in the study to help reduce the confounding influence of scale order such as responder fatigue. This procedure is consistent with Hardre et al.'s (2007) study. The original placement of the instruments was random with the exception of the demographic items being deliberately placed at the end.

Marlowe-Crowne. The Marlowe-Crowne scale (Crowne & Marlowe, 1960) measures an individual's need for approval and consists of 33 true/false questions

about political and social concerns. A sample of Marlowe-Crowne items are (a) It is sometimes hard for me to go on with my work if I am not encouraged, and (b) I sometimes try to get even, rather than forgive and forget. Scores are calculated by adding positive (true) items together, resulting in a score from 0–33. An individual with a high score exhibits a greater need for approval. Crowne and Marlowe first utilized this scale with a sample of 300 college students ($M = 15.5$, $SD = 4.4$). This and other studies reported means ranging from 12.3 to 16.4 and a Cronbach's alpha from .73 to .88 (Fisher, 1967; Paulhus, 1984; Tanaka-Matsumi & Kameoka, 1986). The Marlowe-Crowne scale continues to be used in modern leisure and social research, with similar means and alpha coefficients (Ferrando & Anguiano-Carrasco, 2010; Miotto & Preti, 2008; Ward & Ellis, 2008).

Morally debatable behaviors. The original scale (Harding et al., 1986) consists of 22 statements referring to morally debatable behaviors and measures what people consider being right and wrong, or moral judgment. Only 21 statements were used in this study, as the statement "killing in self-defense" was removed from the scale based on IRB's recommendation. A sample of Morally Debatable Behavior items are (a) Someone accepting a bribe in the course of duties and (b) Failing to report damage you've done accidentally to a parked vehicle. The participant then responds from 1 (*never justified*) to 10 (*always justified*) on a Likert scale. Scoring the Morally Debatable Behaviors scale results in a total score as well as individual scores for three subscales. The subscales are personal-sexual morality (8 items), self-interest morality (8 items), and legal-illegal morality (8 items). Three items are calculated in both self-interest morality and legal-illegal morality. The items' average is taken for the total scale and subscales with a high score representing tolerance of morally debatable behaviors and a low score representing moral strictness, each ranging from 1 to 10. Harding et al. (1986) report a range of total means across 10 European countries from 2.12 to 3.17 and subscale means of 3.53 (personal-sexual morality), 2.28 (self-interest morality), and 1.93 (legal–illegal morality). The Morally Debatable Behavior scale continues to be used in leisure and social research (Begue, 2001; Staats, Hupp, Wallace, & Gresley, 2009).

The way I feel about myself. This scale is used to assess how children and adolescents feel about themselves and consists of 80 yes/no statements (Piers, 1984). Twenty items were dropped from the original scale because they were redundant with other items and it was desired to shorten the length of this study's overall questionnaire. Samples of this scale's items are (a) My classmates make fun of me, and (b) When I grow up, I will be an important person. This scale is scored by summing all *yes* responses for the total scale. The scale consists of six subscales or clusters representing individuals' feelings about self based on behavior (14 items), intellectual and school status (16 items), physical appearance and attributes (11 items), anxiety (14 items), popularity (12 items), and happiness/satisfaction (10 items). Some items fall under multiple clusters, while other items do not correlate with individual subscales. Subscales are scored in the same manner as the total scale, summing all *yes* responses. Higher total and cluster scores represent a positive self-evaluation, with lower scores representing a negative evalu-

ation. Pooled responses of 1,183 public school children from grades 4 through 12 resulted in a mean of 51.84 with a standard deviation of 13.87 (Piers, 1984). Pier's mean is based on 80 items representing 64.8% of items used in this survey. Based on ratio comparison, a projected mean for the shortened scale in this study is 38.88. Reliabilities for subscales reportedly range from .78 to .93 (Wolf, Sklov, Hunter, Webber, & Berenson, 1982), while the overall scale had a Cronbach's alpha of .90 (Piers, 1984). This scale continues to be used in leisure and social research (Cardenal & Fierro, 2003; Mishra, 1992).

Leisure satisfaction measure. This scale measures respondents' level of satisfaction with their leisure and consists of 51 statements on a Likert scale ranging from 1 (*never true*) to 5 (*always true*; Beard & Ragheb, 1980). Samples of Leisure Satisfaction Measure items are (a) I enjoy doing my leisure activities and (b) My leisure activities give me a sense of accomplishment. The responses' mean is calculated, resulting in a total score from 1–5, with scores greater than 4 representing high leisure satisfaction and scores less than 2 representing low leisure satisfaction. Additionally, scores are calculated in the same manner for six subcategories: psychological (13 items), educational (12 items), social (11 items), relaxation (4 items), physiological (6 items), and aesthetic (5 items). The Leisure Satisfaction Measure has good face validity and a high degree of reliability, with an overall Cronbach's alpha of .96 and subscale reliabilities ranging from .85 to .92 (Beard & Ragheb, 1980). Leisure and social researchers continue to use the original scale and shortened versions of the scale (Gerber et al., 2006; Wang, Chen, Lin, & Wang, 2008).

Leisure boredom. This is a 16-item Likert scale designed to measure an individual's perception of his or her leisure participation as boring (Iso-Ahola & Weissinger, 1987). Samples of Leisure Boredom items are (a) Leisure time is boring, and (b) Leisure time activities do not excite me. The scale ranges from 1 (*strongly disagree*) to 5 (*strongly agree*), and scores are summed, resulting in a total score from 16–80, with higher scores representing higher leisure boredom. Iso-Ahola and Weissinger (1987) report a reliability of .90 in a study sample of 400 community residents. Researchers continue to use this scale to measure leisure boredom (Lin, Lin, & Wu, 2009; Wegner, Flisher, Chikobvu, Lombard, & King, 2008).

Satisfaction with life scale. This five-item scale measures an individual's satisfaction with life (Diener et al., 1985). A sample of this scale's item is as follows: In most ways my life is close to ideal. Each item consists of a Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*), determining a participant's agreement with statements referring to life satisfaction. Total summed scores range from 5–35, with a higher score representing more satisfaction with life. In their study of 176 undergraduate students, Diener et al. report a mean of 23.5 (*SD* = 6.43) and a Cronbach's alpha of .87. This instrument continues to be used, along with modified versions, by leisure and social researchers alike (Agate et al., 2009; Johnson, Zabriskie, & Hill, 2006; Wang et al., 2008).

Analysis

Data were entered into SPSS version 18 for the paper/pencil instruments. The online data was imported directly into SPSS from Qualtrics. Both sets of data were then cleaned and analyzed. Scores from the online version and paper/pencil versions of each instrument were compared using a paired sample t-test. Significant differences were examined at the .05 level. If a significant difference between the online and paper/pencil data sets was identified, an effect size (Cohen's d) was calculated for that comparison. For further reference, the demographics collected from the online and the paper/pencil versions were also compared. It was believed demographics were stable and would not change depending on the method used to report them.

In order to further investigate differences that may be more sensitive to data collection techniques, subscales within instruments were also investigated. It was hypothesized the subscales may be more sensitive due to the specific measurement of particular aspects versus the overall construct measured by the scale. The subscales of instruments that were significantly different were compared using paired sample t-tests followed by calculating Cohen's d.

Repeated measures Hierarchical Linear Modeling (HLM) was then used to investigate if there was a relationship between personal characteristics and if the participant responded differently to the data collection methods. This statistical technique was used in order to nest each participant's responses within the individual. HLM is an appropriate method to use for nested data because it "allowed testing the effects of the situational variables that were nested within participants as well as testing the effects of the individual differences variables" (Sibthorp, Witter, Wells, Ellis, & Voelkl, 2004, p. 91). Furthermore, HLM relaxes the assumption of independence of observations.

The HLM analysis first examined the null model for each of the scales with significant differences between paper/pencil and online methods. From the null models an interclass correlation (ICC) was calculated for each scale to determine the proportion of unexplained variance. Next, the variable method was included as a Level 1 fixed effect. Method's slope was fixed because Thum and Bryk (1997) and Ma, Ma, and Bradley (2008) suggest not allowing Level 1 variables slope to be random unless it is the primary focus of the research question. As suggested by Ma et al. (2008), this model, with method's slope being fixed, was not interpreted because it is an intermediate step in the modeling building process, especially since the focus of this model is personal characteristics of the participants. It is important to note, however, that method was significant as a Level 1 contributor for all models. Amount of variance explained by method was also calculated. Next, personal characteristics (gender, year in school, personal income, family income, time spent on computer each week, and technology savvy) were included for each of the scales' models as Level 2 variables. Only significant personal characteristics were retained in the complete model, and the amount of additional variance explained was calculated by comparing the model with only method included as a fixed Level 1 variable with the model of method as a Level 1 variable and personal characteristics of the included as Level 2 variables. Individual subscales were not modeled because the purpose of the study was to gain a general feel for the per-

sonal characteristics influencing different responses between the two methods. The researchers believed these general trends could be found by examining the total scales. Therefore, differences in subscales were only analyzed via paired sample t-test as stated above.

Results

Descriptive characteristics and comparisons between paper/pencil and online responses were reported for each of the six scales (see Table 1). There were no significant differences in scores for the Marlowe-Crowne, Leisure Boredom, and Satisfaction With Life Scales, while the Morally Debatable Behaviors scale, The Way I Feel About Myself scale, and the Leisure Satisfaction Measure did have a significant difference in the way participants responded to the same items on the two different methods for data collection. It is, however, important to consider the effect size (Cohen's d) when interpreting these significant differences. For all of the significant findings, the effect size was small, suggesting the groups are distributed similarly around a similar mean (Salkind, 2005); however, the effect size for Leisure Satisfaction Measure was mildly moderate (Cohen's $d = .20$). This finding suggests the way participants respond to their leisure activities is slightly different between paper/pencil and online data collection techniques. Participants reported being slightly more satisfied with their leisure when responding in the paper/pencil format.

Sample demographics were collected with both paper/pencil and online techniques, and descriptive characteristics and comparisons were reported (see Table 2). These data served as a comparison group because the demographics should not vary from the first to second data collection. The measured demographics were gender, year in school, personal income, family income, ethnicity, marital status, time on computer, and technology savvy. The only demographic to indicate a significant difference between methods was technology savvy ($t = -10.14$, $p < .01$). The demographics considered being the most stable (gender, year in school, ethnicity, and marital status) did not suggest a significant difference in response format.

Instruments' subscales were investigated in order to further determine differences that may be more sensitive to data collection technique. The Morally Debatable Behaviors subscales consisted of personal-sexual morality, self-interest morality, and legal–illegal morality. From the three subscales, participants responded differently on self-interest morality ($M_{pp} = 2.72$, $M_o = 2.45$, $t_{(131)} = 2.70$, $p < .05$) and legal–illegal morality ($M_{pp} = 2.28$, $M_o = 2.08$, $t_{(131)} = 2.47$, $p = .05$) with the paper/pencil responses being higher in terms of representing greater tolerance of morally debatable behaviors. The effect size was considered mildly moderate at .21 and small at .19, respectively (Salkind, 2005).

The Way I Feel About Myself subscales include behavior, intellectual and school status, physical appearance, anxiety, popularity, and happiness/satisfaction. Of these subscales, responses were significantly different for behavior ($M_{pp} = 12.59$, $M_o = 12.29$, $t_{(131)} = 2.37$, $p < .05$), intellectual and school status ($M_{pp} = 11.90$, $M_o = 11.56$, $t_{(131)} = 2.45$, $p < .05$), physical appearance ($M_{pp} = 8.80$, $M_o = 8.45$, $t_{(131)} = 2.33$, $p < .05$), anxiety ($M_{pp} = 9.06$, $M_o = 8.71$, $t_{(131)} = 2.05$, $p < .05$), popularity (M_{pp}

= 8.40, M_o = 8.10, $t_{(131)}$ = 2.08, $p < .05$), and happiness/satisfaction (M_{pp} = 8.48, M_o = 8.08, $t_{(131)}$ = 2.28, $p < .05$), and again the paper/pencil responses were higher, representing a more positive self-evaluation than lower scores. The effect sizes of the difference between the paper/pencil and online method of data collection were .15, .13, .15, .10, .12, and .21, respectively.

The Leisure Satisfaction Measure subscales include psychological, educational, social, relaxation, physiological, and aesthetic. Of these subscales, responses were significantly different for education (M_{pp} = 3.61, M_o = 3.50, $t_{(129)}$ = 2.58, $p < .05$), social (M_{pp} = 3.72, M_o = 3.59, $t_{(129)}$ = 2.75, $p < .05$), physiological (M_{pp} = 3.57, M_o = 3.42, $t_{(129)}$ = 2.61, $p < .05$), and aesthetic (M_{pp} = 3.73, M_o = 3.60, $t_{(129)}$ = 2.28, $p < .05$), and again the paper/pencil responses were higher for each subscale, representing more satisfaction with life than lower scores. The effect sizes of the differences between the data collection methods were .17, .21, .19, and .19, respectively. Participants responded with higher means (being the more socially desirable response) universally on the paper/pencil method.

While the effect sizes are considered small for the subscales and total scales, it was desired to determine if there were personal characteristics trends related to the participants responding differently between the two techniques. This examination was conducted on the scales with significant differences between paper/pencil and online methods (Leisure Satisfaction Measure, The Way I Feel About Myself, and Morally Debatable Behaviors, see Table 3). For the Leisure Satisfaction Measure, the random-coefficient model consisting of method as a Level 1 predictor resulted in method being a significant contributor (B = .09, SE = .02 $t_{(329)}$ = -3.03, $p < .05$). The intraclass correlation (ICC) is .69, suggesting that 31% variance was explained by the Level 1 variable (method), leaving 69% of the variance yet to be explained. When Level 2 variables (personal characteristics) were added to the model, no additional variance was explained between the two data collection techniques, leaving considerable variance yet to be explained (χ^2 = 939.74, $p < .05$).

The Way I Feel About Myself questionnaire had an ICC of .82 with method as the Level 1 predictor (B = -1.55, SE = .40, $t_{(329)}$ = -3.75, $p < .05$). Family income was the only significant Level 2 predictor (B = .58., SE = .16 $t_{(194)}$ = 3.55, $p < .05$). The proportion of increase in explained variance accounted for by adding family income was 6%. The difference between paper/pencil and online responses still had significant variance yet to be explained after Level 2 predictors were accounted for (χ^2 = 1718.29, $p < .05$).

The Morally Debatable Behaviors scale had an ICC of .63 with method entered as a Level 1 predictor (B = -.23, SE = .08, $t_{(329)}$ = -2.82, $p < .05$). Gender (B = -.52, SE = .17, $t_{(193)}$ = -3.09, $p < .05$) and year in school (grand mean centered, B = -.19, SE = .07, $t_{(328)}$ = -2.69, $p < .05$) were the significant Level 2 predictors. The proportion of increase in explained variance accounted for by adding these two variables was 9%. The differences between paper/pencil versus online responses for the Morally Debatable Behaviors scale still had a significant amount of variance to be explained (χ^2 = 683.69, $p < .05$). Considering all Level 2 factors and the three different analyses, there was not an established trend suggesting certain personal characteristics influence online versus paper/pencil responses.

Discussion

The purpose of this study was to examine the differences between the use of online versus paper/pencil data collection techniques. There were no significant differences in scores from the two approaches on the Marlowe-Crowne scale, Leisure Boredom scale, or the Satisfaction With Life Scale. This study takes a step closer to leisure scholars comfortably using either online or paper/pencil versions of these instruments and likely receiving similar results within the environments they are traditionally administered. Whereas, the Morally Debatable Behaviors scale, The Way I Feel About Myself scale, and Leisure Satisfaction Measure all had significant differences between online and paper/pencil versions completed by the same respondent. Researchers must be aware that differences exist on these scales between the paper/pencil and online data collection techniques and how these differences may bias the results. The use of Cohen's d, however, suggested the effect size between the two methods was weak to mild with a range from .16 to .20. This small effect size suggests the magnitude may be minimal at best in actual interpretation of the data.

When these differences were examined further by comparing scores from each method among their subscales, findings indicated participants responded differently to the subscales of self-interest and legal–illegal morality within the Morally Debatable Behaviors scale. The means for both of these subscales were higher for the paper/pencil versions. The means were also higher for The Way I Feel About Myself subscales of behavior, intellectual and school status, physical appearance, anxiety, popularity, and happiness/satisfaction as well as for the Leisure Satisfaction Measure subscales of education, social, physiological, and aesthetic for the paper/pencil version. All means for total instruments and subscales were higher for the paper/pencil versions, suggesting participants tend to offer more socially biased responses for the paper/pencil method of data collection. This consistent pattern may be due to participants feeling a different level of anonymity when responding to sensitive questions in a paper/pencil versus online version of a questionnaire.

It is quite possible participants feel their anonymity is at greater risk with traditional paper/pencil data collection techniques. Cooper (1998) referred to anonymity as one of the three factors that "turbocharge" individuals' behavior on the Internet. He further elaborates that today the Internet offers users a "safer" forum because their identity is anonymous, and users behave in a manner they may not if they believed there was a chance of their identity being known (Cooper et al., 2000). One reason for participants' pattern of responses in this study may be that researchers are viewed as *people* within the process. As such, respondents may offer more socially desirable responses when completing paper surveys because they feel their anonymity is not completely protected. Similar results were found among Internet gamblers in that one of the main reasons they gambled over the Internet was to ensure anonymity; they did not want others to view them in a sociably undesirable way (Wong, 2010).

Furthermore, much of today's population is accustomed to sharing private information (i.e., banking account numbers, medical history) through an electronic

medium. This is demonstrated by comparing how American adults prefer to pay their monthly bills. Sixty-two percent state they would rather use electronic payment instead of paper method to pay bills (eMarketer Digital Intelligence, 2011). The familiarity of sharing private information through using a computer and having it protected may have contributed to participants providing a less socially desirable response via the online format. Participants inherently understand that a paper/pencil survey will at some time be viewed by an individual (i.e., data input, facilitator collecting the questionnaires), possibly leading to respondents' desire to have that individual view them in a positive light, resulting in participants unconsciously completing the survey in a socially biased way. Such results substantiate previous research, suggesting respondents experience less peer pressure and may provide more truthful responses to sensitive questions asked in an online format (Rhodes et al., 2003; Roster et al., 2004; Wright & Schwager, 2008).

The implications of online data collection should be considered for both practitioners and scholars. Practitioners servicing at-risk populations and administering assessments that include sensitive questions may receive responses that are less socially biased by collecting data through an electronic format. Findings from this study also contribute to validating previous scholarship in leisure studies that have used online data collection techniques. According to current findings it must be recognized participants respond slightly more socially biased using the paper/pencil method, but likely not to the extent that would call into question the results of previous works.

Another key point to consider is that in today's technologically advanced world people appear to be comfortable with using a computer and respond similarly to online versus paper/pencil data collection methods. This study suggests researchers clearly have viable options for data collection or they may even combine online and paper/pencil methods when collecting data within the same study. Before simply combining data, however, caution should be taken to compute the proper comparative statistics to ensure no major differences exist between scores from the two methods of data collection.

Although overall findings provide relatively clear support for the use of online data collection in leisure research, possible limitations of this study must be acknowledged. Data for this study were collected from college students who are considered part of the technological generation and are very familiar with interacting through electronic means. Participant responses may be biased toward this age group and might not represent the true computer usage comfort of the general population. Moreover, current college students are part of the millennial generation, known for their individualistic attitudes and high belief in their ability to succeed in life (Strauss & Howe, 1997). This may explain the above average reported mean for The Way I Feel About Myself scale in this study, and this should be taken into consideration when interpreting results of this and similar studies.

The limited number of instruments chosen was an additional limitation for this study. While previous research calls for comparison of all online to paper/pencil instruments, this study was only able to focus on a few of the possible instruments commonly used in leisure research. Results do not suggest online data collection techniques are appropriate for all instruments, but rather the results

contribute to understanding the pattern of differences between the two methods. Furthermore, the findings of this study may have been influenced by characteristics of the scale such as length and response format (Likert scale vs. Yes/No) of the instruments. However, no clear pattern was established in this study based on such criteria.

Areas for future study with regard to online data collection for leisure researchers must consider a wider sample, expand to more instruments, and further investigate what Level 2 variables may influence participant responses. Other possible samples include those who are not as familiar with computer usage and providing confidential information electronically. Additionally, further research should include individuals from various generations, or even clinical samples. Other common instruments should also be examined, particularly in leisure studies, to see how they behave differently through various data collection methods. One major area left for future research is further exploring the influence of personal characteristics (Level 2 variables) on how participants respond to online versus paper/pencil data collection techniques. Understanding how such personal characteristics influence responses will enable leisure scholars to design better studies, particularly when considering data collection method.

References

Agate, J. R., Zabriskie, R. B., Agate, S. T., & Poff, R. (2009). Family leisure satisfaction and satisfaction with family life. *Journal of Leisure Research, 41*(2), 205-223.

Ahern, N. R. (2005). Using the Internet to conduct research. *Nurse Researcher, 13*(2), 55-70.

Alexandris, K., Funk, D. C., & Pritchard, M. (2011). The impact of constraints on motivation, activity attachment, and skier intentions to continue. *Journal of Leisure Research, 43*(1),56-79.

Aluja, A., Rossier, J., & Zuckerman, M. (2007). Equivalence of paper and pencil vs Internet forms of the ZKPQ-50-CC in Spanish and French samples. *Personality and Individual Differences, 43*, 2022-2032.

Bachana, T., Goldberg, L. R., & Johnson, J. A. (1999, November). *WWW personality assessment: Evaluation of an online Five Factor Inventory.* Paper presented at the meeting of Society for Computers in Psychology, Los Angeles, CA.

Ballard, C., & Prine, R. (2002). Citizen perceptions of community policing: Comparing Internet and mail survey responses. *Social Science Computer Review, 20*(4), 485-493.

Beard, J. G., & Ragheb, M. G. (1980). Measuring leisure satisfaction. *Journal of Leisure Research, 12*, 20-33.

Begue, L. (2001). Religious affiliation and social commitment as determinants of moral attitudes in the prevention of AIDS or the fight against poverty. *Journal of Psychology: Interdisciplinary and Applied, 135*(5), 571-576.

Bocarro, J., & Sable, J. (2003). Finding the right P.A.T.H.: Exploring familial relationships and the role of a community TR program in the initial years after a spinal cord injury. *Therapeutic Recreation Journal, 37*(1), 58-72.

Breunig, M. C., O'Connell, T. S., Todd, S., Anderson, L., & Young, A. (2010). The impact of outdoor pursuits on college students' perceived sense of community. *Journal of Leisure Research, 42*(4), 551-572.

Bruton, C. M., Floyd, M. F., Bocarro, J. N., Henderson, K. A., Casper, J. M., & Kanters, M. A. (2011). Physical activity and health partnerships among park and recreation departments in North Carolina. *Journal of Park and Recreation Administration, 29*(2), 55-68.

Buchanan, T. (2002). Online assessment: Desirable or dangerous? *Professional Psychology: Research and Practice, 33*(2), 148-154.

Buchanan, T., Ali, T., Heffernan, T. M., Ling, J., Parrott, A. C., Rodgers, J., & Scholey, A. B. (2005). Nonequivalence of on-line and paper-and-pencil psychological tests: The case of the prospective memory questionnaire. *Behavior Research Methods, 37*(1), 148-154.

Cardenal, V., & Fierro, A. (2003). Factors and correlates of self-concept according to the Piers-Harris scale. *Estudios de Psicologia, 24*(1), 101-111.

Cooper, A. (1998). Sexuality and the Internet: Surfing into the new millennium. *CyberPsychology & Behavior, 1*(2), 181-187.

Cooper, A., Delmonico, D. L., & Burg, R. (2000). Cybersex users, abusers, and compulsives: New findings and implications. *Sexual Addiction & Compulsivity, 7*, 5-29.

Cronk, B. C., & West, J. L. (2002). Personality research on the Internet: A comparison of web-based and traditional instruments in take-home and in-class settings. *Behavior Research Methods, Instruments, and Computers, 34*(2), 177-180.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.

Davis, R. N. (1999). Web-based administration of a personality questionnaire: Comparison with traditional methods. *Behavior Research Methods, Instruments, & Computers, 31*(4), 572-577.

Demmel, R. (2002). Internet addiction: A review. *Sucht: Zeitschrift fur Wissenschaft und Praxis, 48*(1), 29-46.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71-75.

Dillman, D. (2000). *Mail and Internet surveys: The tailored design method* (2nd ed.). New York, NY: John Wiley & Sons.

Dodd, D. C. H., Zabriskie, R. B., Widmer, M. A., & Eggett, P. (2009). Contributions of family leisure to family functioning among families that include children with developmental disabilities. *Journal of Leisure Research, 41*(2), 261-286.

Duffy, M. E. (2002). Methodological issues in web-based research. *Journal of Nursing Scholarship, 34*(1), 83-88.

Ellis, G. D., & Rossman, J. R. (2008). Creating value for participants through experience staging: Parks, recreations, and tourism in the experience industry. *Journal of Park and Recreation Administration, 26*(4), 1-20.

eMarketer Digital Intelligence. (2011). *Electronic vs. paper as payment method preferred by US adult Internet users for paying bills*. Retrieved from http://totalaccess.emarketer.com/Chart.aspx?R=79671&dsNav=Ntk:basic|paper+versus+online|1|,Rpp:50,Ro:26

Ferrando, P. J., & Anguiano-Carrasco, C. (2010). Acquiescence and social desirability as item response determinants: An IRT-based study with the Marlowe-Crowne and the ERQ Lie scales. *Personality & Individual Differences, 48*(5), 596-600.

Fisher, G. (1967). The performance of male prisoners on the Marlowe-Crowne social desirability scale: II differences as a function of race and crime. *Journal of Clinical Psychology, 23*(4), 473-475.

Gallant, K., Smale, B., & Arai, S. (2010). Civic engagement through mandatory community service: Implications of serious leisure. *Journal of Leisure Research, 42*(2), 181-201.

Gerber, L. H., Hoffman, K., Chaudhry, U., Augustine, E., Parks, R., Bernad, M., . . . Mansky, P. (2006). Functional outcomes and life satisfaction in long-term survivors of pediatric sarcomas. *Archives of Physical Medicine and Rehabilitation, 87*(12), 1611-1617.

Gladwell, N. J., Dorwart, C. E., Stone, C. F., & Hammond, C. A. (2010). Importance of and satisfaction with organizational benefits for a multigenerational workforce. *Journal of Park and Recreation Administration, 28*(2), 1-19.

Granello, D. H., & Wheaton, J. E. (2004). Online data collection: Strategies for research. *Journal of Counseling and Development, 82*, 387-393.

Greeff, A. P., & Leroux, M. C. (1999). Parents' and adolescents' perceptions of a strong family. *Psychological Reports, 84*(3), 1219-1224.

Harding, S., Phillips, D., & Fogarty, M. (1986). *Contrasting values in western Europe*. London, England: MacMillan Press.

Hardre, P. L., Crowson, M., Xie, K., & Ly, C. (2007). Testing differential effects of computer-based, web-based, and paper-based administration of questionnaire research instruments. *British Journal of Educational Technology, 38*(1), 5-22.

Hewson, C., & Charlton, J. P. (2005). Measuring health beliefs on the Internet: A comparison of paper and Internet administrations of the multidimensional health locus of control scale. *Behavior Research Methods, 37*(4), 691-702.

Hornberger, L. B., Zabriskie, R. B., & Freeman, P. A. (2010). Contributions of family leisure to family functioning among single-parent families. *Leisure Sciences, 32*(2), 143-161.

Hornig, E. F. (2005). Bringing family back to the park. *Parks & Recreation, 40*(7), 46-50.

Illievia, J., Baron, S., & Healey, N. M. (2002). Online surveys in marketing research: Pros and cons. *International Journal of Market Research, 44*(3), 361-376.

Internet World Stats. (2009, January 16). *United States of America: Internet usage and broadband usage report*. Retrieved from http://www.internetworldstats.com/am/us.htm

Iso-Ahola, S. E., & Weissinger, E. (1987). Leisure and boredom. *Journal of Social and Clinical Psychology, 59*(3), 356-364.

Johnson, A. (1999). Social desirability, anonymity, and Internet-based questionnaires. *Behavior Research Methods, Instruments, & Computers, 31*(3), 433-438.

Johnson, H. A., Zabriskie, R. B., & Hill, B. (2006). The contribution of couple leisure involvement, leisure time, and leisure satisfaction to marital satisfaction. *Marriage & Family Review, 40*(1), 69-91.

Jun, J., & Kyle, G. T. (2011). The effect of identity conflict/facilitation on the experience of constraints to leisure and constraint negotiation. *Journal of Leisure Research, 43*(2), 176-204.

Kelly, J. R. (1996). *Leisure* (3rd ed.). Boston, MA: Allyn & Bacon.

Kimball, A., & Freysinger, V. J. (2003). Leisure, stress, and coping: The sport participation of collegiate student-athletes. *Leisure Sciences, 25*, 115-141.

Klein, J. D., Havens, C. G., & Thomas, R. K. (2009). Comparing adolescent response bias between Internet and telephone surveys. *Journal of Adolescent Health, 44*(2), S36.

Klovning, A., Sandvik, H., & Hunskaar, S. (2009). Web-based survey attracted age-biased sample with more severe illness than paper-based survey. *Journal of Clinical Epidemiology, 62*, 1068-1074.

Lefever, S., Dal, M., & Matthiasdottir, A. (2007). Online data collection in academic research: Advantages and limitations, *British Journal of Educational Technology, 38*(4), 574-582.

Li, X., & Petrick, J. F. (2010). Towards an integrative model of loyalty formation: The role of quality and value. *Leisure Sciences, 32*(3), 201-221.

Lin, C., Lin, S., & Wu, C. (2009). The effects of parental monitoring and leisure boredom on adolescents' Internet addiction. *Adolescence, 44*(176), 993-1004.

Lonsdale, C., Hodge, K., & Rose, E. A. (2006). Pixels vs. paper: Comparing online and traditional survey methods in sports psychology. *Journal of Sport & Exercise Psychology, 28*, 100-108.

Lutner, R. E., Roizen, M. F., Stocking, C. B., Thisted, R. A., Kim, S., Duke, P. C., . . . Cassel, C. K. (1991). The automated interview versus the personal interview. *Anesthesiology, 75*, 394-400.

Ma, X., Ma, L., & Bradley, K. D. (2008). Using multilevel modeling to investigate school effects. In A. O'Connel & D. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 59-110). Charlotte, NC: Information Age.

Mactavish, J. B., & Schleien, S. J. (1998). Playing together growing together: Parents' perspectives on the benefits of family recreation in families that include children with a developmental disability. *Therapeutic Recreation Journal, 32*(3), 207-230.

Mannell, R., & Kleiber, D. (1997). *A social psychology of leisure*. State College, PA: Venture.

McDonald, H., & Adam, S. (2003). A comparison of online and postal data collection methods in marketing research. *Marketing Intelligence and Planning, 21*(2), 85-95.

Miller, E. T., Neal, D. J., Roberts, L. J., Baer, J. S., Cressler, S. O., Metrik, J., & Marlatt, G. A. (2002). Test–retest reliability of alcohol measures: Is there a difference between Internet-based assessment and traditional methods? *Psychology of Addictive Behaviors, 16*(1), 56-63.

Miotto, P., & Preti, A. (2008). Suicide ideation and social desirability among school-aged young people. *Journal of Adolescence, 31*(4), 519-533.

Mishra, K. (1992). A comparative study of achievement motivation and scholastic achievement in relation to self-concept. *Indian Journal of Social Work, 53*(1), 138-142.

Mitas, O., Qian, X. L., Yarnal, C., & Kerstetter, D. (2011). "The fun begins now!": Broadening and building processes in Red Hat Society participation. *Journal of Leisure Research, 43*(1), 30-55.

Mowen, A. J., Payne, L. L., Orsega-Smith, E., & Godbey, G. C. (2009). Assessing the health partnership practices of park and recreation agencies: Findings and implications from a national survey. *Journal of Park and Recreation Administration, 27*(3), 116-131.

Mulvaney, M. (2011). A study of the role of family-friendly employee benefits programs, job attitudes, and self-efficacy among public park and recreation employees. *Journal of Park and Recreation Administration, 29*(1), 58-79.

Nash, J. B. (1953). *Philosophy of recreation and leisure*. Dubuque, IA: Brown.

Orthner, D. K. (1998). Strengthening today's families: A challenge to parks and recreation. *Parks and Recreation, 33*, 87-98.

Orthner, D. K., & Mancini, J. A. (1990). Leisure impacts on family interaction and cohesion. *Journal of Leisure Research, 22*, 125-137.

Palmer, A. A., Freeman, P. A., & Zabriskie, R. B. (2007). Family deepening: A qualitative inquiry into the experience of families who participate in service expeditions. *Journal of Leisure Research, 39*(3), 438-458.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*(93), 598-609.

Pettit, F. A. (2002). A comparison of world-wide web and paper-and-pencil personality questionnaires. *Behavior Research Methods, Instruments, & Computers, 34*(1), 50-54.

Piers, E. V. (1984). *Piers-Harris children's self-concept scale: Revised manual*. Los Angeles, CA: Western Psychological Services.

Poff, R. A., Zabriskie, R. B., & Townsend, J. A. (2010). Modeling family leisure and related family constructs: A national study of U.S. parent and youth perspectives. *Journal of Leisure Research, 42*(3), 365-391.

Raat, H., Mangunkusumo, R. T., Landgraf, J. M., Kloek, G., & Brug, J. (2007). Feasibility, reliability, and validity of adolescent health status measurement by the child health questionnaire child form (CHQ-CF): Internet administration compared with the standard paper version. *Quality of Life Research, 16*, 675-685.

Rademacher, J. D. M., & Lippke, S. (2007). Dynamic online surveys and experiments with the free open-source software dynQuest. *Behavior Research Methods, 39*(3), 415-426.

Rhodes, S. D., Bowie, D. A., & Hergenrather, K. C. (2003). Collecting behavioural data using the World Wide Web: Considerations for researchers. *Journal of Epidemiology and Community Health, 57*(1), 68-73.

Riva, G., Teruzzi, T., & Anolli, L. (2003). The use of the Internet in psychological research: Comparison of online and offline questionnaires. *CyberPsychology & Behavior, 6*(1), 73-80.

Roster, C. A., Rogers, R. D., Albaum, G., & Klein, D. (2004). A comparison of response characteristics from web and telephone surveys. *International Journal of Market Research, 46*(3), 359-373.

Salk, R. J., & Schneider, I. E. (2009). Commitment to learning within a public land management agency: The influence of transformational leadership and organizational culture. *Journal of Park and Recreation Administration, 27*(1), 70-84.

Salkind, N. J. (2005). *Statistics for people who (think they) hate statistics* (2nd ed.). Thousand Oaks, CA: Sage.

Schillewaert, N., & Meulemeester, P. (2005). Comparing response distributions of offline and online data collection. *International Journal of Market Research, 47*(2), 163-178.

Shepherd, R., & Edelmann, R. J. (2005). Reasons for Internet use and social anxiety. *Personality and Individual Differences, 39*, 949-958.

Sibthorp, J., Witter, E., Wells, M., Ellis, G., & Voelkl, J. (2004). Hierarchical linear modeling in parks, recreation, and tourism research. *Journal of Leisure Research, 36*(1), 89-100.

Staats, S., Hupp, J. M., Wallace, H., & Gresley, J. (2009). Heroes don't cheat: An examination of academic dishonesty and students' views on why professors don't report cheating. *Ethics & Behavior, 19*(3), 171-183.

Strauss, W., & Howe, N. (1997). *The fourth turning: An American prophecy*. New York, NY: Broadway Books.

Swinton, A. T., Freeman, P. A., Zabriskie, R. B., & Fields, P. J. (2008). Nonresident fathers' family leisure patterns during parenting time with their children. *Fathering: A Journal of Theory, Research, & Practice About Men as Fathers, 6*(3), 205-225.

Tanaka-Matsumi, J., & Kameoka, V. A. (1986). Reliabilities and concurrent validities of popular self-report measures of depression, anxiety, and social desirability. *Journal of Consulting and Clinical Psychology, 54*(3), 328-333.

Thum, Y. M., & Bryk, A. S. (1997). Value-added productivity indicators: The Dallas system. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 100-109). Thousand Oaks, CA: Corwin.

Topp, N. W., & Pawloski, B. (2002). Online data collection. *Journal of Science Education and Technology, 11*(2), 173-178.

Touvier, M., Mejean, C., Kesse-Guyot, E., Pollet, C., Malon, A., Castetbon, K., & Hercberg, S. (2010). Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. *European Journal of Epidemiology, 25*, 287-296.

Tu, H., Chen, L., Wang, M., & Lin, J. (2007). The impact of neuroticism on leisure satisfaction: Online game application. *Social Behavior and Personality, 35*(10), 1399-1404.

Wang, E. S., Chen, L. S., Lin, J. V., & Wang, M. C. (2008). The relationship between leisure satisfaction and life satisfaction of adolescents concerning online games. *Adolescence, 43*(169), 177-184.

Wang, X., & Walker, G. J. (2011). The effect of face concerns on university students' leisure travel: A cross-cultural comparison. *Journal of Leisure Research, 43*(1), 133-147.

Wang, Y., Lee, C., Lew-Ting, C., Hsiao, C., Chen, D., & Chen, W. (2005). Survey of substance use among high school students in Taipei: Web-based questionnaire versus paper-and-pencil questionnaire. *Journal of Adolescent Health, 37*, 289-295.

Ward, P. J., & Ellis, G. D. (2008). Characteristics of youth leadership that influence adolescent peers to follow. *Journal of Park and Recreation Administration, 26*(2), 78-94.

Wegner, L., Flisher, A. J., Chikobvu, P., Lombard, C., & King, G. (2008). Leisure boredom and high school dropout in Cape Town, South Africa. *Journal of Adolescence, 31*(3), 421-431.

Wolf, T. M., Sklov, M. C., Hunter, S. M., Webber, L. S., & Berenson, G. S. (1982). Factor analytic study of the Piers-Harris children's self-concept scale. *Journal of Personality Assessment, 46*(5), 511-513.

Wong, I. L. K. (2010). Internet gambling: A school-based survey among Macau students. *Social Behavior and Personality, 38*(3), 365-372.

Woolhouse, L., & Myers, S. (1999, September). *Factors affecting sample make-up: Results for an Internet-based personality questionnaire.* Paper presented at the 1999 British Psychological Social Society Psychology Section Conference, Lancaster, UK.

Wright, D., Aquilino, W., Supple, A. (1998). A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a survey on smoking, alcohol and drug use. *Public Opinion Quarterly, 62*(3), 331-353.

Wright, B., & Schwager, P. H. (2008). Online survey research: Can response factors be improved? *Journal of Internet Commerce, 7*(2), 253-269.

Young, K. S. (2005). An empirical examination of client attitudes towards online counseling. *CyberPsychology & Behavior, 8*(2), 172-177.

Zabriskie, R. B. (2001). Family recreation: How can we make a difference? *Parks and Recreation, 36*(10), 30-42.