VARIATIONS IN ONLINE

SURVEY DESIGNS

By

CELINA ANN MENDOZA

Bachelor of Arts in Psychology
University of Central Oklahoma
Edmond, Oklahoma
2005

Master of Arts in Psychology
University of Central Oklahoma
Edmond, Oklahoma
2007

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2014

VARIATIONS IN ONLINE

SURVEY DESIGNS

Dissertation Approved:

Dale R. Fuqua

Dissertation Adviser

Katye M. Perry

Mwarumba Mwavita

C. Robert Davis

ACKNOWLEDGEMENTS

"For I know the plans I have for you, says the Lord. They are plans for good and not for disaster, to give you a future, and a hope." New Living Translation.

First, and foremost, I would like to thank my Lord and Savior, Jesus Christ. For without Him, none of this would have been possible.

To my parents, Mom, thank you for always being there for me. Your enduring love, support, encouragement is something I will always treasure. Dad, thank you for always believing in me and for teaching me it's okay not to be perfect. I can only hope I've made you both proud. Love you two will all my heart.

To my sister, Maria: Thank you for always fighting for me even when fighting with each other. You've always looked out for me and for that, I thank you. Love you, sis!

To my nephew, Benton, and my niece, Caelyn: I love you both more than words can express. I pray for a bright future for the both of you. Dream big and never give up on your goals.

To the rest of my family, near and far, on Earth and in Heaven: I love and miss every one of you. Thank you for being a part of this long journey.

A very special thank you to my advisor and dissertation chair, Dr. Dale Fuqua, for his unwavering support and guidance throughout the past four years. Dr. Fuqua, your extensive knowledge and expertise have been invaluable to me. A thousand thanks for being my mentor and educator. I can only pray that one day that I will contribute to the world as you have.
To Dr. Katye Perry: Thank you for your insightful contributions to this project. I have greatly respected and appreciated your instruction as a REMS faculty member and supervisor. I am proud to have been one of your doctoral students. I wish you nothing but the best in the years to come. Happy Retirement!
To Dr. Mwarumba Mwavita: Thank you for supporting and encouraging me throughout this journey. I consider myself blessed to have you as an advocate for my academic and professional careers.
To Dr. Robert Davis: Thank you for serving on my dissertation committee. I am very grateful for your time, support, and encouragement on this project.
Also, to my fellow REMS colleagues and friends: A huge thank you to each of you for your assistance on this project. Your contributions are greatly appreciated and highly respected.

Although not mentioned by name, I would like to thank all of my wonderful friends: Thank you for having confidence in me. Your positive thoughts, warm wishes, and earnest prayers have not gone unnoticed. I am extremely grateful for all of the support and encouragement I received from each one of you. Thank you for being a shoulder to cry on and someone to laugh with. Let the celebration begin!

Name: CELINA ANN MENDOZA

Date of Degree: MAY, 2014

Title of Study: VARIATIONS IN ONLINE SURVEY DESIGNS

Major Field: RESEARCH, EVALUATION, MEASUREMENT, AND STATISTICS

The purpose of the current study was to extend the literature on the use of survey design elements in the construction of online surveys. More specifically, this study aimed to assess the effects of online survey design characteristics on participants' responses to online questionnaires presented in different visual survey designs.

The 216 participants in this study were Oklahoma State University (OSU) students, enrolled at the OSU-Stillwater campus during the Fall 2013 semester. Sixteen individual survey testing conditions were created in which the survey content was identical but the visual design characteristics varied. These conditions were designed with the purpose of revealing statistically significant differences between survey conditions.

The four independent variables used in the study were selected based on their relatively prevalent use in online survey design. Participants responded to questions on three instruments (i.e., knowledge test, attitudinal questionnaire, and computer self-efficacy scale). Response rate and completion time were also measured. Regardless of the visual presentation, all participants viewed the same questionnaire items (i.e., content).

A series of four 2x2x2x2 completely randomized factorial (CRF) analysis of variances (ANOVAs) were performed in order to determine the effects of the independent variables on the dependent measures. Qualitative analyses were also performed and data involved identifying and analyzing participants' text responses to an invitation to provide feedback regarding their experience during the survey.

This study implored a mixed mode strategy in which different forms (N=16) of the same online survey were administered to different groups. In general, participants revealed a strong dislike for the survey colors selected for this study. Taken together, the results of this study provide considerable evidence for the influence of color in online survey design. While the present study explored the effects of color on participants responses to questionnaire items presented in an online format, the unanticipated effects of the participants' physical environments necessitates future investigation. Future studies should bear in mind that surveys should be designed to be aesthetically pleasing and incorporate color in a way that makes the survey interesting while producing salient, quality data.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Surveys play a central role in data collection in nearly every research oriented discipline (Miller, 2001) including, but not limited to, applied social science, education, health, and marketing and business (Shapiro, Bessette, Baumlin, Ragin, & Richardson, 2004). The behavioral scientist may use survey techniques to provide useful insights into the cognitive and behavioral processes comprising human interaction with the surrounding environment (Krosnick, 1999). Educational organizations and government agencies (Gannon, 1973) often use surveys to aid in the construction of a program, policy, or procedure (Middlestadt, Bhattacharyya, Rosenbaum, Fishbein, & Shepherd, 1974). Marketing firms and other consumer-oriented agencies often rely on psychological concepts to develop surveys evaluating a particular product or service (Prunk, 1994). By measuring respondents' self-reported values, attitudes, and behaviors, survey researchers can explore or describe relationships between variables (Passmore, Dobbie, Parchman, Tysinger, 2002). Therefore, data collection through use of survey methods provides the basis for many types of research attempting to explain, assess, or understand people's viewpoints and perceptions toward the research construct or domain.

The emergence of the Internet has all but replaced paper-based communication in most workplaces. Reliance on computers has significantly increased over the past 10-15 years, obliging companies and other agencies to employ telecommunication methods. Likewise, survey methodology is continuing to embrace technological advances (Kim & Huynh, 2008) and has considerable advantages over paper-and-pencil administration modes of data collection (Oppermann, 1995; O'Rourke, 2011). Such advantages include cost, time, labor (Birnbaum, 2004), response validation (Shapiro, Bessette, Baumlin, Ragin, & Richardson, 2004), and a higher degree of anonymity (Howell, Rodzon, Kurai, & Sanchez, 2010) and confidentiality. The fast and efficient nature of online survey design software has made web-based, or Internet, survey research more appealing, yet simple to use. This is especially important for researchers in academic institutions where efficiency is highly valued. Surveys administered via the Internet allow researchers to invite large numbers of individuals to participate in a research study. Often, researchers will generate a random sample of participants from a general target population such as college students (Mitra, Jain-Shukla, Robbins, Champion, & DuRant, 2008; Sax, Gilmartin, & Bryant, 2003) to include in the study.

In their review of the literature, Howell, Rodzon, Kurai, and Sanchez (2010) found inconsistent evidence for the comparability of computer-based Internet questionnaires to paper-and-pencil questionnaires. Past research (Brock, Barry, Lawrence, Dey, & Rolffs, 2010; de Beuckelaer & Lievens, 2008; Perkins, 2004) has shown data collected via the Internet may not produce the same factor structure or other psychometric properties (i.e., mean scores, reliability, etc.) as data collected through an equivalent paper version. Furthermore, Brock, Barry, Lawrence, Dey, and Rolffs (2010) found two main differences between administration modes that can affect data integrity (i.e., accuracy of responses, completion of survey, response rate, etc.). Such differences included environmental factors (e.g., social distractions) and clinical factors (e.g., anonymity). In their review of the literature, Barenboym, Wurm, and Cano (2010) found inconsistency regarding the effect of the testing environment on task performance. More

precisely, while all of the studies claimed to use the same questionnaire for both modes of data collection, testing conditions (i.e., environment) varied.

Sax, Gilmartin, and Bryant (2003) contended that when two or more modes are presented either sequentially or concurrently, administration modes must produce equivalent outcomes. Researchers often employ multiple modes in order to combat the limitations (such as small sample size) associated with using a single administration mode. However, mixed mode designs are susceptible to certain mode effects making the use of multiple survey modes within a single administration impractical (Lugtig, Lensvelt-Mulders, Frerichs, & Greven, 2011). Using different modes, such as a postal mail survey and a web-based survey, may impact how participants respond and thereby results in nonresponse biases. If researchers elect to implement a mixed-mode design, they should do so with careful consideration to how the survey is developed, designed, and implemented with respect to the survey itself and the sampling procedures.

Fan and Yan (2010) described the process of designing a web survey as comprised of four steps: 1) web survey development, 2) web survey delivery, 3) web survey completion, and 4) web survey return. Stylistic features are included in every step of this process and play a key role in the overall appearance of the web survey (Couper, 2000). Brock, Barry, Lawrence, Dey, and Rolffs (2010) found visual presentations of online surveys can be interpreted very uniquely with respect to the individual viewer. In contrast to paper-based methods, web surveys often require more cognitive processing on the part of the participants. Specifically, visual displays (i.e., computer settings such as screen size and resolution) may influence how participants interpret items on a questionnaire which subsequently influences data integrity. Furthermore, stylistic design elements such as font color, background color, and other visual enhancements have shown to negatively impact the quality of responses (Stern, Smyth, & Mendez, 2012). Researchers who wish to utilize web surveys should carefully consider their effects on the participants' ability to perceive computerized items in the same manner as paper-based methods.

The increasing dependency on computers and technology to communicate information has forced the general population to become familiar, if not proficient, with computer-based programs such as the Internet and the various software programs involved. Data collection via online methods must rely on those who have access to the Internet (Perkins, 2004) and are able to properly respond to requests for participation. Although the increase in technology has allowed researchers more opportunities to recruit participants through various telecommunication methods, the population in which results can be generalized to is limited by the study's "coverage area" (i.e., population of Internet users; Howell, Rodzon, Kurai, & Sanchez, 2010; Kwak & Radler, 2002; Schmidt, 1997). Due to lack of Internet availability, certain demographic groups are often excluded in web-based survey research studies and therefore cannot be included in the representativeness of the coverage area. Regardless of who is included in the study, the ability of one to effectively use a computer and interact with the various software platforms is crucial.

Online software tools for survey development enable researchers to employ computerized self-administered questionnaires in their research studies. Businesses, government agencies, and non-profit organizations rely on these tools to gather data from consumers, employees, or the public. When deciding on which software to use, companies must consider the breadth of the research design, including the number of participants to be sampled. Researchers must also determine the procedures that will be used to develop and distribute the survey as well as collect and analyze responses. The intricacies of the questionnaire's visual and response formats must also be considered. While numerous software programs exist to meet the various demands of the research, very few offer free complete access to their services. Most are limited by the number of questions that can be included in a single survey and regulate survey distribution to a certain number of participants. These programs offer basic services and are often devoid of certain visual design features and response format options.

Qualtrics, a research software company that provides researchers in any field the ability to create highly sophisticated surveys, is a simple yet attractive web-based survey software

package that offers its users many design features that are unavailable in similar programs. The software is especially attractive to academic researchers because institutions with licenses to Qualtrics are able to provide faculty, staff, and students with free access to all of the survey software's features. Such features include advanced options for question types (i.e., response options, response validation, branching, skipping, etc.) and the ability to integrate visual design features such as background color and text color with design elements such as a question highlight bar. Taken together, Qualtrics is a user-friendly software program that provides researchers multiple options to effectively and methodically customize a web-based survey.

In order to assess one's capability to use a computer, self-efficacy must be considered. Several measures of computer self-efficacy exist in accordance with the various aspects of self-efficacy theory (Conrad & Munro, 2008). Participants' reported level of computer experience, comfort, and skill, as well as time spent engaged in technology, are all dimensions of self-efficacy that influence attitudes towards computers (Compeau & Higgins, 1995). According to Torkzadeh and van Dyke (2002), those who lack knowledge of computer technology may encounter some anxiety and avoid participation in difficult computer-based tasks, such as online surveys. Conversely, highly efficacious computer users may be more capable of navigating numerous Internet-based programs due to their familiarity with various platforms. Furthermore, individuals who have encountered these platforms may not be influenced by design elements that often distress less-accustomed Internet users. Taken together, the effects of highly sophisticated, intricately designed online surveys have not been considered in relation to an individual's judgment about their ability to meet the situational demands associated with online surveys (Moos & Azevedo, 2009).

Statement of the Problem

According to several researchers (Howell, Rodzon, Kurai, & Sanchez, 2010; Kwak & Radler, 2002; Vallejo, Mañanes, Comeche, & Diaz, 2008), modes of administration are only comparable if questionnaires are presented as similarly as possible. Previous research has shown

inconsistent findings regarding the appropriateness for comparing alternative modes (e.g., paper-based and web-based), yet little is known about the standards for making strict comparisons among multiple online survey formats. Specifically, the various design elements available for constructing online surveys are extensive with regard to stylistic elements such as font color, background color, and combinations thereof. The visual display of surveys is especially important in understanding how participants perceive and respond to items on a questionnaire. For example, the appearance of such stylistic elements such as a poorly contrasted text to background color combination has been shown to negatively impact the validity of the data. Collectively, the many differences in how textual information can be presented in an online format have not been thoroughly investigated.

Purpose of the Study

The ease and appeal of the Internet to conduct survey research has allowed researchers to consider the various ways in which surveys and questionnaires can be visually presented to participants. Specifically, the visual effects in web-based surveys can be varied numerous ways including, but not limited to, using assorted colors for font and background, selecting assorted typefaces for text, and applying other stylistic features. Although previous literature has suggested that online surveys should be designed so they resemble their paper-based counterparts as closely as possible, this may not necessarily be true for comparing multiple online survey formats (i.e., multiple administrations of the same questionnaire presented in different design formats). The effects of varying the visual design of multiple online surveys have not been extensively researched. Therefore, the purpose of this study is to extend the literature on the effects of variations in the visual presentation of online survey designs.

Significance of the Study

This research could contribute valuable information in regards to the appropriateness of varying design characteristics of online surveys. More precisely, this research contributes to our understanding of the effects of these characteristics on how participants respond to items on a

questionnaire. The use of color in online presentation of materials including the interaction of text and background color is especially important in understanding how information appears visually to respondents, especially in an online format (Garcia & Caldera, 1996). Additionally, the presence or absence of visual aids (i.e., progress bar and highlight bar) to indicate a participant's current position in the survey process may contribute to our understanding of the impact of these features in online survey responses.

The potential finding that participants' responses differ in regards to text color and background color and whether these choices either aid or diminish both the aesthetic design of an online questionnaire as well as the legibility and readability of text (in accordance with a contrasting background) would represent the appropriateness of using visual enhanced designs to improve response quality (Deutskens, Ruyter, Wetzels, & Oosterveld, 2004). Moreover, the impact of either including or excluding a progress bar and/or highlight bar in the contextual composition, or visual display, of the online survey must be assessed in order to understand how participants are influenced by visual cues. Specifically, the interaction of any of these design features involve psychological responses to stimuli (i.e., presented information), which may affect performance and lead to measurement biases.

Definition of Terms

For the purposes of this study, the following terms are defined as such:

Completion time: the time it takes for participants to respond to individual items or total survey completion time.

Computer self-efficacy: an individual's self-reported judgment about their ability to use a computer.

Measurement effects: the influence of the survey on a participant's responses to items on an instrument. Essentially, responses are is influenced by the mode in which the respondent participates and occurs when respondents respond to survey questions differently depending on the survey mode (Howell, Rodzon, Kurai, & Sanchez, 2010).

Mixed mode*:* a strategy in which two or more survey modes are combined to collect data (Lugtig, Lensvelt-Mulders, Frerichs, & Greven, 2011). Vannieuwenhuyze and Loosveldt (2012) defined a mixed mode design as procedures in which data from different respondents are collected by different data collection modes.

Mode effect: occurs when the mode of data collection influences responses to questionnaire items. There are two types: measurement effects and selection effects.

Online (or Internet or web-based) survey: a conventional form of data collection in which participants are recruited through web-based techniques, such as email (Miller, 2001). Researchers typically have no direct contact with the participants.

Paper-and-pencil administration: mode of delivery in which a researcher collects data through printed materials such as via postal mail or in face-to-face settings (Hardré, Crowson, & Xie, 2012).

Paper (or paper-based) survey: method of data collection that involves printed materials. This mode of data collection includes a paper and pencil questionnaire wherein participants manually respond to survey items (Fowler, 2009, p. 69).

Questionnaire: research tool comprised of several items (i.e., questions) administered to participants using paper-based or web-based methods for the purpose of gathering self-reported information from respondents (Webb, 2000). Questionnaires are generally self-administered wherein participants respond to questions without the presence or aid of the researcher.

Response rate: can refer to the rate in which participants respond to and complete the survey questionnaire.

Selection effects: differences in respondent characteristics. Selection effects occur when participants choose to respond to one mode over another (Vannieuwenhuyze, Loosveldt, & Molenberghs, 2010).

Survey methodology: method for collecting information (i.e., data) about a particular construct, or concept, in a research domain. Methods for collecting data are specific to the aims of the research

(Couper, 2000). These methods generally consist of an instrument, or series of instruments, comprised of several items related to the investigation's research questions.

Survey mode (or administration mode): method in which questionnaire is administered (Fricker & Schonlau, 2002). Mode refers to the method of survey delivery (i.e., paper-based or web-based) and the design characteristics related to that particular survey (Kwak & Radler, 2002).

Web-based administration: mode of delivery in which data is collected via the Internet, typically through an online software program (Alessi & Martin, 2010).

Research Questions

The following research questions will guide this study:

1. Does text color, background color, appearance of progress bar, or appearance of highlight bar affect the a) response rate, b) completion time, and group scores on c) a knowledge test, d) attitudinal measure, and e) self-efficacy scale when administered in a web-based mode?
2. Are there significant interactions of text color, background color, appearance of progress bar, or appearance of highlight bar on a) response rate, b) completion time, and group scores on c) a knowledge test, d) attitudinal measure, and e) self-efficacy scale when administered in a web-based mode?

CHAPTER II

REVIEW OF THE LITERTURE

The purpose of this chapter is to provide a review of the literature relevant to the research proposed in this document. The first section presents a brief introduction to the purposes, processes, and contributions of survey research. The second section describes the major modes of administration currently used in survey methodology. This section is divided into two subsections: 1) paper-based surveys and 2) online surveys. The advantages and disadvantages for both types of administration modes are also provided. Section three includes a thorough, yet succinct description of the impact of contextual effects and visual design elements in online, or web-based, surveys. The fourth section compares conventional methods (i.e., paper-and-pencil measurement) to more contemporary survey modes (i.e., Internet-administered). More specifically, section four explains how responses may be influenced by the type of survey mode employed. Section five describes the integration of survey modes and how data quality is affected by the utilization of multiple survey modes. This section also explores the methodological issues present in using a mixed mode strategy. Section six provides a brief explanation of the different types of measurement error commonly found in survey research.

The last section, reviews the contribution of respondents' familiarity with computer technology and its relationship to survey methodology. This chapter ends with a summary of the major findings of prior research, limitations of these findings, and presents the potential contributions of the current study to the existing body of literature.

Survey research

Researchers often employ survey methods—research tools known as questionnaires (Tourangeau, 2004)—to gather evidence, or data, from samples of people (designated as participants or respondents) in order to infer "perceptual, cognitive, and affective experiences and processes" (*see* Karabenick et al., 2007) of the survey respondents. Questionnaires are generally self-administered, self-report measures, or instruments in which researchers collect a diverse yet specific type of information including, but not limited to: 1) demographics, 2) personal histories, 3) knowledge, 4) attitudes, and 5) behaviors (Passmore, Dobbie, Parchman, & Tysinger, 2002). According to Passmore, Dobbie, Parchman, and Tysinger (2002), surveys are either exploratory (i.e., researchers explore cause-and-effect or correlational relationships between variables) or descriptive (i.e., participants report attitudinal or behavioral information about themselves). Passmore and colleagues also asserted surveys can either address one (i.e., knowledge) or more (i.e., behavior and attitude) underlying construct in a single questionnaire. Thus, an instrument can either consist of a unidimensional scale measuring a single construct or multidimensional scale measuring two or more constructs.

In a review of recent research topics (i.e., research articles published between 2000 and 2010) published in *Public Opinion Quarterly*, an interdisciplinary journal providing a wide range of social science research, topics included, but were not limited to, "investigations of methodological issues involved in survey validity including questionnaire construction, interviewing and interviewers, sampling strategy, and mode of administration" (Public Opinion Quarterly, 1937-present). Schaeffer and Dykema (2011) examined the recent contributions of survey methodology and summarized their findings according to various issues in measurement

(e.g., question development, context, wording, format, form, or implementation including mode of administration). The authors also explored the role of "topic" in how participants respond to items on a questionnaire. Topic, according to Schaeffer and Dykema, can refer to either a broad general domain such as 'health' or a more specific question focused on a certain aspect of content domain (i.e., comparison of one's health to others'). Regardless of the topic (i.e. research questions), researchers must consider the level of cognitive processing, attention, and any other possible personal and situational factors needed to complete the survey and how these factors may influence measurement (Lietz, 2010; Schaeffer and Dykema; 2011).

Modes of administration

In a general sense, survey mode (or response mode) is the method in which a survey is administered (Fricker & Schonlau, 2002). More precisely, mode refers to the method of survey delivery (e.g., web/Internet, paper and pencil, etc.) and the design characteristics related to that particular survey (Kwak & Radler, 2002). According to Webb (2000), questionnaires are designed specific to the research conditions. The construction of the survey, and subsequent administration of the survey, depends on the aims of the research project (Couper, 2000). Thus, many factors must be considered when choosing the mode of administration (Passmore, Dobbie, Parchman, & Tysinger, 2002). The type of information that will be collected, the sample of participants that the information will be collected from, the types of data analysis that are to be used, and budget and time constraints are the principal bases from which a questionnaire is designed (Passmore, Dobbie, Parchman, & Tysinger, 2002; Webb, 2000).

When comparing modes of survey administration, researchers often examine the quality of the data as determined by response patterns such as response rate and response speed (Fricker & Schonlau, 2002). Bowling (2005) identified survey response rate (including total survey completion and responses to individual questionnaire items), researcher bias, and sampling including the characteristics of the respondents or accuracy of answers (e.g., inability to answer, unwillingness to answer, etc.; Webb, 2000) may occur as a result of using human participants for

data collection. These sources of error can be broadly summarized into two categories: measurement errors and non-measurement errors (Bowling, 2005). Measurement errors occur as a result of the construction and implementation of the survey instrument and procedures used in the data collection process. Relatedly, non-measurement errors are artifacts of survey design, sampling methods, and response rate. How participants respond to different survey modes has been repeatedly investigated with no consensus to why results occur as they do. Previous research has shown little consensus on the exact effects of administration mode on participants responses (Carini, Hayek, Kuh, Kennedy, & Ouimet, 2003).

In a meta-analysis on the effects of mode (i.e., questionnaire administration) on data collection, Bowling (2005) found inconsistent or inconclusive results. While some studies indicated a strong effect of survey mode on participants' responses, others revealed little to no influence of questionnaire administration. Differences in the methods of contact (or sampling) and other systematic design features made any strict comparison between studies challenging. That is, because surveys are conducted in regards to a specific research design, any discrepancies in data quality could be attributable to any of the design characteristics of the research endeavor (Fan & Yan, 2010; Roster, Rogers, Hozier, Baker, Albaum, 2007). In a review of the literature on the impact of administration mode, Schwarz, Strack, Hippler, and Bishop (1991) found response effects in survey measurement vary as a function of survey mode as well as how surveys are designed.

Paper and pencil surveys

Paper and pencil, or paper-based, surveys are one of the most traditional modes of data collection (Wood, Nosko, Desmarais, Ross, & Irvine, 2006). Hardré, Crowson, and Xie (2012) described paper-based questionnaires as printed "hardcopy" materials that are individually handed out to each research participant and returned to the researcher after survey completion (Fowler, 2009, p. 69). Therefore, paper-based assessments are self-administered self-reports in which participants respond to set of questions measuring some construct. This type of survey

methodology can be accomplished through several strategies. Researchers can send questionnaires through postal mail to a mass quantity and variety of participants or in face-to-face settings where a sample of respondents is gathered and a questionnaire is distributed in a group environment.

Before the Internet emerged as the primary survey method of data collection, paper-and pencil surveys were the most economical and efficient alternative to more labor-intensive interviewing techniques (Wood, Nosko, Desmarais, Ross, & Irvine, 2006). Primarily, paper based surveys can include a vast variety of participant samples based upon the specific needs of the project. A researcher may purposefully sample a select group of participants (e.g., college students, government employees, consumers, etc.; Mitra, Jain-Shukla, Robbins, Champion, & DuRant, 2008) or include a vast variety of participants to include in the study. Similarly, paper based surveys are predominantly administered in group settings. (Fowler, 2009, p. 69) and allow for, in some cases, a mass collection of data in a single setting (Brannon, 1981). Perkins (2004) and Sax, Gilmartin, and Bryant (2003) found a higher rate of completion (i.e., response rate) using this method.

Stanton (1998) provided many benefits of administering a paper-and-pencil questionnaire such as the control of samples representative of the general population. Although these methods are applicable to web-based surveys as well, paper-based surveys differ in that they are perhaps the only method that allows researchers to gather information from populations not easily or readily accessible (e.g., persons with limited financial resources, lower educated individuals, older people, etc.; Vicente & Reis, 2010). Similarly, participants do not need to be computer literate or have experience using Web procedures for paper-based surveys. Additionally, paper and pencil survey administration can offer more protection against participants' privacy (Brannon, 1981). For example, researchers can control for both confidentiality and anonymity by removing any identifying information such as participants' names, email addresses, and other unique identifiers (Alessi & Martin, 2010; Perkins, 2004) from paper questionnaires. In paper-

based administration modes, in most cases, the researcher has control over the testing environment or condition. Unlike web survey methods, participants complete the questionnaire under procedures (such as time, location, etc.) strictly set by the researcher. Also dissimilar to web-based surveys, paper and pencil administration is associated with lower procedural problems during data collection. Namely, paper-based surveys are not as susceptible to computer or software malfunctions as with Internet surveys.

Disadvantages. Ryan, Corry, Attewell, and Smithson (2002) found several problems arose when participants responded to a paper version of their questionnaire. First, the researchers observed where participants gave more than one response to a single item question. They also found paper versions allowed for participants' responses to be ambiguous; that is, paper versions allowed participants to place marks on the line between two overlapping categories and inside instruction boxes. These issues are not often present in web survey modes as the researcher can use special techniques to restrict multiple responses to each of the questionnaire's items and force participants to respond to every item (Pouwer, Snoek, van der Ploeg, Heine, & Brand, 1998; Ryan, Corry, Attewell, & Smithson, 2002). Kwak and Radler (2002) contended total survey completion or low item nonresponse is indicative of good survey quality. In paper-and-pencil administered questionnaires, there is no guarantee or requirement that participants respond to every item. Additionally, most paper-based questionnaires are to a certain degree constrained by how individual survey items are presented to respondent. Few variations are available in how questionnaires are formatted, displayed and presented to respondents (Fowler, 2009, p. 69).

One of the major drawbacks of paper-and-pencil administration survey methods includes the amount of labor involved regarding the data. While the collection of data may occur very easily and quickly, the transmission of that data into a software analysis program may prove difficult and time-consuming. In other words, paper-based survey methods do not allow for data to be entered directly into a system or catalogue of data. That is, a researcher must manually enter the data into a database. The manual entry of data into a database is more prone to a large error

rate due on part of the person entering the data (i.e., keying errors; Shapiro, Bessette, Baumlin, Ragin, & Richardson, 2004). Careful time and consideration must be taken to ensure the data is accurate. If inaccuracies are present, additional time is necessary to correct these inaccuracies (Ryan, Corry, Attewell, & Smithson, 2002). Furthermore, administering a paper survey potentially requires more time and effort on the parts of both the researcher and the participants. Possible sources of labor include the process of setting up appointment times and places for the researcher and participant or participants to meet and conduct the survey. Taken together, the limitations of paper-based surveys draw attention to the need for more feasible and efficient modes of data collection, such as surveys administered over the Internet.

Web-based surveys

Before the popularity of the Internet emerged in the 1990s, data collection via the use of questionnaires was limited to more traditional survey techniques such as postal mail, random-digit dialing (RDD) telephone, and face-to-face (FTF) administration (Couper, 2011; Dillman & Christian, 2005). To combat issues associated with these techniques, software developments in the 1990s made it even more possible, and more appealing, for researchers to conduct surveys using online methods (Oppermann, 1995; O'Rourke, 2011). The arrival of the 21st century brought with it the rapid development of the Internet (also called the World Wide Web or the Web) as a contemporary tool for researchers to collect data from a large population of individuals (Richman, Weisband, Kiesler, & Drasgow, 1999; Cook, Heath, & Thompson, 2000; Couper & Miller, 2008; Schmidt, 1997). The number of households and persons who have access to the Internet has continued to increase over the years allowing for more potential respondents to be involved in research studies (Gaddis, 1998). Howell, Rodzon, Kurai, and Sanchez (2010) advocated for the use of web surveys due to their ability to target specific populations, such as college students (Mitra, Jain-Shukla, Robbins, Champion, & DuRant, 2008; Sax, Gilmartin, & Bryant, 2003). Due to the various characteristics of web-based surveys, they are broadly defined

by Mitra, Jain-Shukla, Robbins, Champion, and DuRant (2008) as an "Internet-based approach in which participants are electronically sent email links directing them to the online questionnaires."

Similar to the traditional paper and pencil format of data collection, web-based surveys are self-administered (Fricker & Schonlau, 2002; Mitra, Jain-Shukla, Robbins, Champion, & DuRant, 2008). However, unlike their more traditional counterparts, the administration of a web survey generally occurs at a time and place of the participant's convenience (Malhotra, 2008; Sax, Gilmartin, & Bryant, 2003). Due to the nature of the Internet, web surveys are available 24 hours a day provided the respondent has an active Internet connection (Perkins, 2004; Vicente & Reis, 2010). Depending of the aims of the research design, respondents have a limitless amount of time to complete the survey and can return to the survey at a later time if he or she does not complete the questionnaire in a single session (Cook, Heath, & Thompson, 2000; Mitra, Jain-Shukla, Robbins, Champion, & DuRant, 2008). Researchers can also program surveys so that respondents can review and verify their answers for accuracy (Christian, Dillman, & Smyth, 2007; Vallejo, Mañanes, Comeche, & Diaz, 2008). This procedure aids the researcher by ensuring all responses are appropriate for submission into the pre-established database (this advantage is discussed more fully below in the *Advantages: Labor* section).

In terms of participants' rights, computerized or web-based assessments allow for a great degree of anonymity (Howell, Rodzon, Kurai, & Sanchez, 2010) while responses to standard paper-and-pencil versions may not fully guarantee the anonymity, or confidentiality, of the respondent's answers. Thus, the online questionnaire is thought to provide more valid responses. This is particularly true for self-report measures in which participants provide ratings about themselves (e.g., participants may give socially desirable answers; Passmore, Dobbie, Parchman, & Tysinger, 2002) or questionnaires involving sensitive information (Lewis, Watson, & White, 2009). Passmore, Dobbie, Parchman, and Tysinger (2002) found these tendencies may bias responses thereby threatening the validity of the survey measurement.

As stated before, unlike the more conventional modes of data collection, web surveys can be programmed to include the validation of responses (Shapiro, Bessette, Baumlin, Ragin, & Richardson, 2004). That is, the survey can be designed so that respondents are required to answer every single item on the survey questionnaire (Gwaltney, Shields, & Shiffman, 2008; Kwak & Radler, 2002) resulting in a lower nonresponse rate (Stanton, 1998). Christian, Dillman, and Smyth (2007) stressed this strategy allow respondents to check for errors in responding such as inaccurate answers or missed responses to questions. This strategy all but guarantees total survey completion. In a comparison between paper-based and web-based surveys, Denscombe (2006) found a higher completion rate for the web-based questionnaire option showing support for the claim that electronic surveys administered via the Internet produce higher completion rates than their paper-based counterparts. Additionally, web-based or Internet surveys are generated so that a mass collection of data can gathered in a relatively short amount of time (Mitra, Jain-Shukla, Robbins, Champion, & DuRant, 2008).

Advantages. Web surveys are increasingly gaining considerable advantages over more conventional modes of data collection (Kaplowitz, Hadlock, & Levine, 2004) and are increasingly used as the predominant mode of data collection in surveying methods (Couper & Miller, 2008). Cost, time, and labor are the most commonly cited motives for researchers who choose a web survey over the more traditional alternatives or who included a web survey in the design methodology (Barenboym, Wurm, & Cano, 2010; Fricker & Schonlau, 2002; Weible & Wallace, 1998; Vicente & Reis, 2010). Additionally, variations of web based survey methods displayed certain advantages over visual designs of paper-based administration modes (Brock, Barry, Lawrence, Dey, & Rolffs, 2010). Thus, due to their feasibility and ease of operation and implementation, web surveys have the potential to all but replace traditional methods of data collection (Barenboym, Wurm, & Cano, 2010; Couper, 2001; Oppermann, 1995; Richman, Weisband, Kiesler, & Drasgow, 1999).

Cost. Using the Internet to conduct research and collect data via web surveys is associated with lower cost (Cook, Heath, & Thompson, 2000; Schmidt, 1997). Couper (2001) and others (Cook, Heath, & Thompson, 2000; Cobanoglu, Warde, & Moreo, 2001) found using the Internet web-survey data collection method saved publishing, distribution, and subsequent survey collection costs (Schmidt, 1997) such as postage stamp, data entry, and printing expenses (Kwak & Radler, 2002; Mitra, Jain-Shukla, Robbins, Champion, & DuRant, 2008).

Time. Several researchers (Pouwer, Snoek, van der Ploeg, Heine, & Brand, 1998; Schmidt, 1997) reported the ease of administering a web survey includes the ability to reach a large number of participants. Likewise, the development of personal computers or the availability of a computer with Internet access has provided an abundance of potential respondents the opportunity to participate in online web surveys. Couper (2001) suggested this method not only allows for a larger population to access the survey but responses to the survey can occur in less time than it would take to mail out survey instruments. Additionally, response speed, or the time required for a survey to be returned, is typically faster for web surveys (Oppermann, 1995). The immediacy of the Internet also allows for researchers to gather data more efficiently (Kwak & Radler, 2002).

Labor. Pouwer, Snoek, van der Ploeg, Heine, and Brand (1998) and Ryan, Corry, Attewell, and Smithson (2002) showed support for claims that web-based studies are more efficient in terms of the time it takes to enter questionnaire data into a database. Most software programs allow for the automatic entry of the participants' responses into a database (Alessi & Martin, 2010; Schmidt, 1997). The automatic nature of the computerized version of a survey avoids error on the part of the researcher manually collecting, entering, and verifying the data into the database (Kwak & Radler, 2002; Richman, Weisband, Kiesler, & Drasgow, 1999).

Taken together, past research has shown considerable advantages of using the Internet to conduct survey research. Compared to paper-and-pencil administration modes of data collection, web-based questionnaires are believed to be more efficient in terms of cost, time, labor, and allow

researchers more design alternatives (*discussed below*). Still, despite their increasing appeal, web-based survey methods suffer from limitations that researchers should thoroughly considered when choosing the mode of administration.

Disadvantages. Although the Internet is believed to provide an efficient and effective way to collect data from a large population of individuals, it is limited to the population of Internet users (Howell, Rodzon, Kurai, & Sanchez, 2010; Schmidt, 1997). Kwak and Radler (2002) referred to this as "coverage area." Although widespread use of the Internet provides more opportunity to reach a diverse sample of respondents, Howell, Rodzon, Kurai, and Sanchez (2010) found web surveys to be biased. For example, participants who are older, less educated, and have lower socioeconomic statuses are less likely to use the Internet on a regular basis (Stanton, 1998) and be unable to discern the various computer software platforms (e.g., Qualtrics, Survey Monkey, etc.; Gay, Mills, & Airasian, 2012, p. 194) used for online data collection. Thus, online survey methods to collect data are regulated by the availability of Internet users and their results cannot be representative of non-Internet users (Roster, Rogers, Hozier, Baker, & Albaum, 2007).

Although computer technology has increased exponentially since its advent, there remain certain demographic groups that are less prone or less able to use Internet communication (Shapiro, Bessette, Baumlin, Ragin, & Richardson, 2004). Therefore, one of the main shortcomings of web surveys is their tendency to have purposive or probability-based (Couper, 2000) samples in which a specific population is targeted for participation in a research study (Howell, Rodzon, Kurai, and Sanchez, 2010; Stanton, 1998). This technique limits the representativeness of the general population, or the generalizability of results (Buchanan & Smith, 1999; Couper, 2000; Curasi, 2001; Sax, Gilmartin, & Bryant, 2003; Schmidt, 1997). Couper (2000) identified this as error, an attribute resulting from the study's coverage area—"a function of the mismatch between the target population and the frame population"—which prevents generalizability.

Curasi (2001) argued the online method of data collection differs significantly from other methods in that prospective participants are often solicited through online techniques such as email (Fricker & Schonlau, 2002). Several problems may arise when soliciting participants in this manner such as recipients simply ignoring and deleting an email requesting their participation and emails returned to the sender as "undeliverable" (Bowling, 2005; Sax, Gilmartin, & Bryant, 2003; Ward, Clark, Zabriskie, & Morris, 2012). There is also no guarantee recruited participants check their email on a regular basis potentially resulting in a lower response rate. Populations also exist in which individuals are less Internet-savvy or are concerned about the privacy of their responses, particularly if the survey questionnaire requires the admission of sensitive personal information, potentially increasing dropout or nonresponse rate (Bowling, 2005; Curasi, 2001). Sax, Gilmartin, and Bryant (2003) found low response may also be due to a general disinterest in research participation or in the research topic.

Other disadvantages characteristics of online surveys include biases resulting from features of both the participants and the survey mode (Heerwegh & Loosveldt, 2008). Howell, Rodzon, Kurai, and Sanchez (2010) discussed the possible effects of the environment on participants' responses to an online questionnaire. The conditions in which a participant completes a web survey are unknown and therefore uncontrollable by the researcher. Given that participants are authorized to complete the survey at their own leisure, there is no consistency in testing condition or environmental factors (Brock, Barry, Lawrence, Dey, & Rolffs, 2010). Hardré, Crowson, and Xie (2012) referred to these variations as "potential effects of context beyond the system itself," and involve asocial and social distractions that allow for biases in participants' responses. Characteristics of the participants such as a participant's general mentality (or affect; Hardré, Crowson, & Xie, 2012) at the time of survey completion lead to responses bias which consequently threatens the validity of the data. Participants' location when completing the survey, including conditions under which participants complete the survey and how participants' respond to the survey, influence the quality of data in online survey assessment

(Brock, Barry, Lawrence, Dey, & Rolffs, 2010). Such characteristics are specific to each participant and cannot be observed or measured by the researcher. Thus, there is a lack of standardization in Internet administered surveys (Miller, 2001).

Survey design: Context and visual effects in web-based surveys

Brock, Barry, Lawrence, Dey, and Rolffs (2010) and Couper, Traugott, and Lamias (2001) found differences between paper-and-pencil survey modes and how information is presented online may impact how participants respond to questions on a survey. Several investigations (Gaddis, 1998; Singh, Taneja, & Mangalaraj, 2009) recommended researchers consider testing the web survey on various computers and monitors in order to better understand the diversity of the participants' experiences viewing the online questionnaire. For example, the visual presentation of a survey on the Web may require different strategies for interpreting items because of the various computer processors (e.g., Windows, Mac, etc.) and visual displays (i.e., computer screen size, screen settings, and browser settings; Brock, Barry, Lawrence, Dey, & Rolffs, 2010). Due to the various features (e.g., size, style, and color of font) of online survey software programs, researchers also have the ability to design surveys in a variety of survey modes (Couper, Traugott, & Lamias, 2001; O'Rourke, 2011; Walston, Lissitz, & Rudner, 2006). However, varying these elements within Internet administration has shown to impact the integrity of the data.

Richman, Weisband, Kiesler, and Drasgow (1999) cited computerized web-based questionnaires that have been converted from a paper-and-pencil administered version. If transferred correctly (*see section on Equivalence below*), a web survey is simply a paper-based mode of data collection that has been transitioned into an online format (Sethuraman, Kerin, & Cron, 2005). Singh, Taneja, and Mangalaraj (2009) recommended researchers should carefully consider the type or types of response options (i.e., Likert scale, pull-down menu, etc.) available and allow for any possible responses to all questions. Gwaltney, Shields, and Shiffman (2008) furthered the argument for the equivalence of paper-and-pencil versions to their computerized

counterparts based on the scores from the respondents given the response options and item content were the same. Daley, McDermott, Brown, and Kittleson (2003) suggested that in order to test the suitability of a paper-and-pencil survey for use on the Internet, the survey should first be conducted in a paper-based format. Passmore, Dobbie, Parchman, and Tysinger (2002) stressed that during this initial phase, or pilot test, researchers should gather feedback on the readability of the questions (i.e., how questions are worded), survey flow (i.e., order of questions), and other stylistic features of the questionnaire (e.g., typeface, font size, color, etc.).

Another concern for researchers involves the effects of survey design, or how items on a questionnaire are developed, organized, and presented to the respondent. Couper, Tourangeau, and Kenyo (2004) cautioned survey researchers to consider the potential issues or effects of survey design on how questions may be interpreted by respondents. To combat unintended measurement errors (*described below*), Fan and Yan (2010) suggested survey researchers follow basic principles when writing items for use on a questionnaire. These principles include how items are constructed in terms of their linguistic properties. Specifically, items should be structured so that they 1) use simple language, 2) are without bias, and 3) avoid ambiguity. For web-based questionnaires, Couper (2000) found three basic components of web survey development: 1) how questions are worded, 2) the order in which questions are presented, and 3) the visual design (*described below*) of the questions. Any one of these has the potential to influence measurement error (Schaeffer & Dykema, 2011; Schwarz, Strack, Hippler, & Bishop, 1991; Shropshire, Hawdon, & Witte, 2009).

An additional, more recent advantage of online administration includes an increase in the development of software platforms. Advancing technology has provided researchers more opportunities to construct an online survey using certain visual effects such as color and font (Singh, Taneja, & Mangalaraj, 2009). Ling and van Schaik (2006) found variations in the construction of a web-based survey include, but are not limited to, stylistic elements such as fonts, colors, and spacing. Additionally, visual enhancements such as photo images (Shropshire,

Hawdon, & Witte, 2009) and video (Couper, Tourangeau, & Kenyon, 2004) are believed to influence participants' responses in questions in a web-based survey. Regarding early phases of web survey design, Daley, McDermott, Brown, and Kittleson (2003) found participants reacted negatively to a pretest of a Web survey using stylistic designs typically seen in paper-based surveys (i.e., 12-point Times New Roman black font on a white background). Their finding further advocated for the increase of the overall attractiveness of web-based surveys using graphics or visual images, non-traditional typefaces, and colors.

Fan and Yan (2010) found the Web offers numerous alternatives for survey design through visual enhancements such as verbal and visual elements (Couper, Tourangeau, & Kenyon, 2004). While verbal information includes question wording, survey instructions, and response formats, visual information refers to the appearance of questions on the page (typeface including color and size, background color, and graphics used as prompts or directives for survey completion). Gwaltney, Shields, and Shiffman (2008) also stressed the importance of evaluating the effects of questionnaire format and design. Several studies have shown minor variations in the design of survey items tended to influence participants' responses (Brock, Barry, Lawrence, Dey, & Rolffs, 2010; Stern, Smyth, & Mendez, 2012). More specifically, Best and Krueger (2004) suggested using contrasting colors—light text on dark background —to assist in the respondent's ability to read the information on the screen. According to Hill and Scharff (1999), the legibility of survey questions is best when a dark colored text is presented on a light colored background.

When conducting a web-based survey, Ling and van Schaik (2002) observed six main technical issues researchers associated with the construction of questionnaires: 1) survey design, 2) appearance, 3) scrolling versus paging, 4) context effects, 5) progress indicators, and 6) testing conditions. Such features of web survey design have shown to influence or contribute to measurement error. Task elements—how questions are worded, response options, instructions or navigational cues, etc.—include visual elements and are often confounded by stylistic elements, or the overall layout (i.e., font, color, and background color, etc.) of the Website or questionnaire

(Couper, Tourangeau, & Kenyon, 2004). The survey's overall appearance (survey characteristics such as layout, font, and color) and construction (including question wording and flow of the questionnaire) both lead to biases which compromise the validity of the data. Decisions about the visual design regarding which colors (for both font and background) and fonts (including size and type) to use are governed by the researcher's preference but should be carefully executed. For example, researchers should avoid using too many colors (Ling & van Schaik, 2002), consider separating individual questions with thin lines, and choose a standard font that can be easily read by the participant.

Researchers may also consider using visual illustrations or images in web survey designs. More features of web designs concerns the number of questions presented per page. While no evidence exist recommending one design over the other, research has shown advantages in using "session bars" or progress indicators (Conrad, Couper, & Tourangeau, 2003; Couper, Traugott, & Lamias, 2001; Vicente & Reis, 2010) as a visual cue for participants to assess their current progress in the survey. Conrad, Couper, and Tourangeau (2003) explored the similarity of paper-based questionnaires to web administered surveys in relation to their ability to communicate information about respondents' progress. In paper surveys, participants can estimate their completion time based on the number of pages remaining in the survey packet or booklet. For web-based questionnaires, respondents may be informed of their current progress by means of a scroll bar which can be designed to serve one of two purposes. Namely, a progress bar can indicate the percentage of the survey that has been completed or how much of the survey remains (Vicente & Reis, 2010).

Ling and van Schaik (2002) and Schmidt, Liu, and Sridharan (2009) found the use of color in visual displays is related to the cognitive and perceptual abilities of its viewer. Researchers or designers of online surveys should not only base decisions about what colors to use on individual preference, but also on the "aesthetic qualities of particular colors or color combinations". For example, Humar, Gradišar, and Turk (2008) found the effects of color

combinations on visual performance have shown some combinations are perceived more favorably than others. Thus, the use of color in online survey construction should be carefully selected and implemented accordingly. Web survey designers should consider using contrasting colors (e.g., black text on white background, blue text on yellow background, red text on green background, etc.) to help facilitate performance on measures (Ling & van Schaik, 2002). Humar, Gradišar, and Turk (2008) related color to readability and legibility of online text. Both of these features are psychological responses to stimuli (i.e., presented information) which may affect performance.

The reliance of text in web-based studies has led researchers to investigate the effects of variations in typefaces when constructing online surveys. In their review of the literature on the usability of the Web, Ling and van Schaik (2006) found a number of variables must be considered when displaying content over the Internet. These "human factors" play a vital role in how questions in an online survey are perceived and answered (Ling & van Schaik, 2007). Researchers should design surveys and present information in a way that is visually appealing to and readable by the respondent. Specifically, researchers should consider using typeface, spacing, and color for both aesthetic and pragmatic purposes (Ling & van Schaik, 2007). In Ling and van Schaik (2007), previous research suggested web page designers avoid scrolling (i.e., breaking up large amounts of text into several pages) and present text in a highly readable format (e.g., consider typeface including font size, spacing, and color). Ling and van Schaik referred to these as typography variables which have been shown to affect response time and accuracy of performance on tasks.

Vicente and Reis (2010) acknowledged the possible effects of the "visual language" expressed by online surveys. Such language includes, but is not limited to, font size, font type, color, layout, and graphic images (Couper, 2001). Dillman, Tortora, Conradt, and Bowker (1998) also studied the effects of using "basic" or a "sophisticated" visual design in online surveys and found non-response rate is lower for surveys with more elaborate features. Vicente and Reis

advised researchers should carefully consider the use of visual enhancements (such as color, font type, and font size) in online survey design. Specifically, questionnaires should be purposefully designed to visually engage respondents' interest—in terms of the survey's overall aesthetics and appeal (Couper, Traugott, & Lamias, 2001; Oliver, 2002; Vicente & Reis, 2010)—but avoid unnecessary or potentially distracting design features (Couper, Traugott, & Lamias, 2001; Schmidt, Liu, & Sridharan, 2009).

Taken together, the differences in how paper-based and Internet surveys are administered (i.e., contextual and visual effects) make it difficult to discern the equivalence of the two methods. Couper, Traugott, & Lamias (2001) and Webster and Compeau (1996) suggested inconsistent findings are due, in part, to how the information (i.e., questionnaire) is presented to participants. That is, the format of the questionnaire is different between survey modes. In their review of the research, Webster and Compeau found these format differences consisted of characteristics regulated by the aims of the researcher (i.e., number of items presented on each page, validation of answers, etc.) as well as visual design effects such as variations in typeface (including the color and size of the font) and background page color.

Comparing survey modes

Several researchers (e.g., Lewis, Watson, & White, 2009; Weigold, Weigold, & Russell, 2013) found various approaches have been used when assessing the comparability of Internet questionnaires to their paper-based counterparts. In a comparison of a computerized version to a standard paper and pencil version of an equivalent questionnaire, Ryan, Corry, Attewell, and Smithson (2002) reported the standards for accepting the equivalence of a computerized (web-based) questionnaire to conventional paper versions. In their review of the literature, Ryan and colleagues described the American Psychological Association's (APA) strict guidelines for the comparison of electronic and paper versions of similarly structured questionnaires. According to the APA,

> "scores on conventional and computer administrations may be considered equivalent when (a) the rank order of scores of individuals tested in alternative modes closely approximate each other, and (b) the means, dispersions, and shapes of the score distributions are approximately the same, or have been made approximately the same by rescaling the scores from the computer mode."

A consensus of the literature reviewing the equivalency of paper-based and web-based surveys indicated the utility of a web survey is only appropriate when it is comparable to other traditional survey methods (Howell, Rodzon, Kurai, & Sanchez, 2010; Kwak & Radler, 2002). Vallejo, Mañanes, Comeche, and Diaz (2008) furthered the argument for equivalent questionnaires as compared by their psychometric properties (factor structure, reliability, mean ratings, etc; Brock, Barry, Lawrence, Dey, & Rolffs, 2010; de Beuckelaer & Lievens, 2008; Perkins, 2004) and suggested online questionnaires be presented as similar to paper-and-pencil surveys as possible, including its design layout (*described above*; Webster & Compeau, 1996). Results from a large-scale, multi-national experiment assessing the measurement equivalence of paper-and-pencil and Internet survey modes showed empirical support for the comparability of paper-based and web-based questionnaires according to their scalar equivalence (i.e., factor loadings; de Beuckelaer and Lievens, 2008). Brock, Barry, Lawrence, Dey, and Rolffs (2010) stressed the importance of interpreting scores (i.e., equal means, variances, and covariances) from computerized versions of conventional paper-and-pencil assessments. They specified that before a questionnaire is administered on the Internet, it should first be "independently evaluated for psychometric equivalence" in a paper-based form.

Gwaltney, Shields, and Shiffman (2008) found differences in how items are presented may influence how a participant responds to a computerized assessment of those items. They also identified the potential difficulties participants may have in understanding the intricacies of an online assessment. When assessing differences in scores between survey modes, Gwaltney, Shields, and Shiffman advised researchers to carefully consider the factors influencing these differences. Participants may respond differently due to a number of factors including, but not

limited to, changes in 1) the respondent's attitude or feelings about the survey topic or individual questionnaire items, 2) mood, condition, or environment in which the survey is completed, or 3) simple random error. Both paper-and-pencil and Internet administration of questionnaires varies widely based on a number of environmental and clinical factors which influences the comparability of paper-and-pencil and Internet survey methods (Brock, Barry, Lawrence, Dey, and Rolffs, 2010).

Few studies have explored the influence of survey mode on the type or quality of information collected, specifically involving the direct comparison of responses to web-based and paper-based questionnaires. To study the effects of mode on data collection, researchers will authenticate a participant's responses by requiring the completion of two survey modes and comparing the results from the measurements. While some studies (Carlbring, Brunt, Bohman, Austin, Richards, Ost, & Andersson, 2007) have shown nearly identical responses from respondents who received either a paper-and-pencil questionnaire or questionnaire administered via the Internet, others (Lewis, Watson, & White, 2009) found different modes of delivery did not produce highly similar data. In their study on methodological implications for using the Internet to collect data, Bertot and McClure (1996) found certain precautions must be taken to insure the consistency between printed (i.e., paper and pencil) and electronic (e.g., Internet) survey. For example, the presentation format and delivery of a traditional paper and pencil questionnaire is inherently different than that of an online survey (Brock, Barry, Lawrence, Dey, & Rolffs, 2010).

Weigold, Weigold, and Russell (2013) found previous research has failed to consider the effects of methodological and statistical issues that prevent the comparability of different survey modes. Specifically, the authors addressed the problems underlying the equivalence of paper-and-pencil measures to Internet, or web-based, survey administration modes. In their review of the literature examining the comparability of paper-and-pencil survey methods to Internet modes of survey administration, Weigold, Weigold, and Russell (2013) attributed the inconsistent findings to differences in administration mode, methodological or statistical approaches, or both.

Controlling for both methodological (i.e., recruitment procedures, assignment of participants to survey modes, and data collection procedural conditions) and statistical effects (i.e., data equivalence and analyses), the researchers found adequate support for the equivalence of paper-and-pencil and Internet survey modes. Still, future research should choose the best possible and most appropriate mode of survey administration based on the aims of the research (Couper, 2000; Medway & Fulton, 2012; Weigold, Weigold, & Russell, 2013).

Taken together, the results of studies assessing comparability of paper-based and web-based surveys indicate the importance of understanding how participants respond to a web administered questionnaire as compared to the standard pencil and paper version of the same (or equivalent) questionnaire. Any differences, significant or slight, in scores on equivalent questionnaires may be due, in part, not only to any or all of the aforementioned factors but also carryover effects or consequences of the participants' familiarity with survey questionnaire items. Participants may simply repeat their answers from the previously administered questionnaire, potentially leading to inaccurate or misappropriated data. Also of issue is participants' tendency to give socially desirable answers when responding in modes other than via the computer (Morgan & Harmon, 2001).

Mixed mode methodology

In order to combat the limitations (i.e., data quality) of using a single survey administration mode, several researchers have proposed a mixed mode strategy (Denscombe, 2006; Kaplowitz, Hadlock, & Levine, 2004; Lugtig, Lensvelt-Mulders, Frerichs, & Greven, 2011; Vannieuwenhuyze, Loosveldt, & Molenberghs, 2010). This technique involves collecting data from different participants (Vannieuwenhuyze and Loosveldt, 2012) using multiple survey modes (e.g., mail surveys, web surveys, telephone surveys) based upon the aims of the researcher (Medway & Fulton, 2012). Some studies employ an alternative mode strategy in which one of at least two equivalent forms of a questionnaire is randomly assigned to participants. Alternatively, participants may be required to complete the same questionnaire more than once although

through different modes of administration. Employing this strategy may also counterbalance the disadvantages of using a single mode administration with the advantages of another survey mode (Jäckle, Roberts, & Lynn, 2010; Medway & Fulton, 2012).

Another feature of mixed mode designs includes the ability of researchers to present surveys either sequentially or concurrently. Sequential mixed mode strategies consist of presenting participants an opportunity to respond to one of two survey modes during the initial phase of data collection then providing non-responders to participate in the second survey mode (Heerwegh, 2009). In contrast, concurrent mixed mode strategies offer both modes simultaneously to participants. In a meta-analysis to estimate the effect of concurrent Web and mail survey modes, Medway and Fulton (2012) found offering a Web option in mail surveys does not increase the overall rate of response.

The results of this study indicated that if a mixed mode method is chosen, researchers must pay careful attention to both the characteristics of the survey (or the survey design; *discussed above*) and the equivalence (*discussed above*) of the questionnaire in order to diminish erroneous differences that may occur. That is, variations inherent in the different types of survey modes may reduce the comparability of data due to the differences in how both online and paper surveys are designed (Weigold, Weigold, & Russell, 2013). Differences in coverage area, nonresponse, and measurement errors are possible pitfalls of mixed mode designs (Jäckle, Roberts, & Lynn, 2010). These differences may result in mode effects which occur when participants respond differently due to the mode in which the survey is administered (Lugtig, Lensvelt-Mulders, Frerichs, & Greven, 2011).

Mode effects

Employing a mixed mode strategy does not always assure higher data quality and consequently may generate forms of bias called mode effects (Vannieuwenhuyze & Loosveldt, 2012; Vannieuwenhuyze, Loosveldt, & Molenberghs, 2010). Mode effects refer to occurrences where a particular survey administration mode causes different data to be collected (Denscombe,

2006) and is comprised of two forms: selection effects and measurement effects. Vannieuwenhuyze and Loosveldt (2012) defined selection effects as differences in respondent characteristics while measurement effects are best explained by the influence of characteristics of the survey mode. Few studies have tried to separate selection effects from measurement effects. According to Vannieuwenhuyze, Loosveldt, and Molenberghs (2010), selection effects and measurement effects are confounded. Any differences or similarities of responses could be due to the either characteristics of the participants or characteristics of the survey mode. In an attempt to differentiate selection effects from measurement effects, Vannieuwenhuyze and Loosveldt (2012) found insufficient evidence to disentangle the two forms.

With the increasing utilization of mixed mode designs to increase response rate comes a diverse range of methodologies employing several different modes of data collection (Millar & Dillman, 2011). Researchers develop their studies using a combination of web-based, postal mail, telephone, and paper versions of equivalent questionnaires in order to gather enough data to increase effect size and produce generalizability (Vannieuwenhuyze & Loosveldt, 2012). The choice of which mode, or combination of modes, to use should depend upon the research questions or purposes and any restrictions (i.e., cost, time, etc.) present. That is, researchers should choose the "optimal method" of data collection based on the aims of the research (Medway & Fulton, 2012). Thus, to employ a mixed mode strategy is burdened by both its advantages and shortcomings (O'Rourke, 2011). Specifically, mixed mode designs are prone to mode effects but are increasingly in survey research due to their overall feasibility and appeal (e.g., cost effectiveness, timeliness, and potential improvements in data quality; Medway & Fulton, 2012).

The majority of these mixed mode studies employ a randomized design in which participants respond to one of the possible survey mode choices (Howell, Rodzon, Kurai, & Sanchez, 2010; Kwak & Radler, 2002; Sax, Gilmartin, & Bryant, 2003) and are matched on certain demographic characteristics of the participants (Lugtig, Lensvelt-Mulders, Frerichs, &

Greven, 2011). If any differences between responses on administration modes are detected, researchers are apt to conclude mode effects are present. However, mode effects are often confounded with sampling procedures and any differences in outcomes must be interpreted with caution. In a review of studies utilizing a mixed mode approach, several have included using different samples of subjects (and recruitment strategies) to complete a single survey mode and making comparisons between survey modes. Other survey methods involved the use of the same subjects completing at least two survey modes (e.g., online and in person; Carlbring, Brunt, Bohman, Austin, Richards, Ost, & Andersson, 2007). For these studies, participation involved responding to the questionnaire in different testing environments or settings and the scores on the questionnaire are compared. In cases where a multiple modes are utilized for data collection, significant findings are difficult to interpret due to confounding mode effects.

Also important is the strategy in which participants are recruited. Previous research has employed various methodologies for designing a mixed mode or alternative mode method —or methodologies in which two different survey modes were utilized—for both random- and within-subjects design.

Sampling frame. Recruitment of subjects is different depending on the mode or method of data collection (Barenboym, Wurm, and Cano, 2010; Curasi, 2001). The ease of access to the Internet allows for greater administration of web surveys to a broad range of individuals. One specific group of individual who possibly have greater access to the Internet includes college students whose communication with instructors, classmates, and administrative staff relies heavily on the World Wide Web (e.g., email, online courses, etc.; Jones, Johnson-Yale, Millermaier, & Pérez, 2008). Most, if not all, universities offer every student a free email address and access to the Internet via computer labs. Researchers can randomly sample a group of students to send a virtual (web) survey or a hard copy of the equivalent questionnaire to another randomly sampled group of students (Miller, 2001). Kaplowitz, Hadlock, and Levine (2004) utilized the randomized, matched method as have numerous other researchers (Barenboym,

Wurm, and Cano, 2010; Kwak & Radler, 2002; Vereecken & Maes, 2006). However, sampling using this method (random, matched) does not allow for strictly comparable samples (Kwak & Radler, 2002).

Ryan, Corry, Attewell, and Smithson (2002) also stressed the importance of matching samples. The responses of participants who complete multiple questionnaire modes must be matched in order to assure data quality and completeness. However, matched samples cannot be fully assured or assessed in studies that utilize a randomized design. In their study, Barenboym, Wurm, and Cano (2010) administered a questionnaire in two distinct yet equivalent methods. Participants had the option to complete the survey in a personal online environment (i.e., at a time and location of their convenience using a computer of their choosing) or in a university computer laboratory setting wherein the time and location were predetermined and set up by the researchers. The latter method included participants using a university-owned computer located in the university's psychology computer lab to complete the questionnaire. Results showed participants' ratings differed as a function of testing condition. One of the major limitations of this study concerned using two different samples of subjects for the two surveys modes. The researchers were unable to conclude whether the significant findings were due to mode effects (characteristics of the survey mode or testing conditions) or person effects (using different participants for the two survey modes).

Measurement error

Both online surveys and paper-and-pencil questionnaires are prone to nonresponse bias (Curasi, 2001; Heerwegh & Loosveldt, 2008) which occurs when participants respond differently than those who did not respond (Sax, Gilmartin, & Bryant, 2003). Although online surveys can be designed so that participants are required respond to each item on the questionnaire, this strategy may discourage participation and lead to dropout. Hence, web based surveys are threatened by their ability to gather the same information as other survey methods. Similarly, just as in paper-and-pencil administration, participation in online surveys can lead to responses bias—or ways in

which participants respond to the items on a questionnaire—which also threatens the validity of the data (Bowling, 2005). Sax, Gilmartin, and Bryant (2003) reviewed several types of response bias that may occur in both online and paper survey administration. Examples of errors include social desirability, acquiescence (i.e., repeatedly endorsing items regardless of content), satisficing (i.e., expending little effort in the interpretation and answering of questions; Heerwegh & Loosveldt, 2008), and exaggeration (or overclaiming) are potential influences that threaten the validity of the data.

The use of computers and the Internet to collect data from participants is becoming increasingly evident. Similar to paper-and-pencil administration, web-based survey research employs self-report instruments, or questionnaires, wherein participants respond to some construct, or content domain. Assuming the two methods are equivalent, web-based instruments are prone to the same measurement issues present in paper-based instruments (Booth-Kewley, Larson, & Miyoshi, 2007). One such issue includes the impact of social desirability. Barenboym, Wurm, and Cano (2010) acknowledged the possibility of participants giving socially desirable answers when in the presence or proximity of the researcher or others (e.g., an in-person survey administered in a group setting). Participants who believe they are being observed, either indirectly or directly, may behave more consciously than they would if completing the questionnaire in an environment of their choosing (e.g. in the privacy of their own home using their personal computer; Howell, Rodzon, Kurai, & Sanchez, 2010). Joinson, Woodley, and Reips (2007) reported lower levels of socially desirable responses in surveys and questionnaires administered via the Internet rather than paper-and-pencil modes due to the high degree of anonymity web-based survey research can provide (Joinson, 1999). This is especially true for survey research exploring personal, sensitive information (Booth-Kewley, Larson, & Miyoshi, 2007).

Another threat to the validity of questionnaires occurs when participants are given limited response options, such as Yes/No, Agree/Disagree, and True/False. This form of bias is known as

acquiescence, the tendency to endorse or refute items regardless of their content (Krosnick, 1999). Acquiescence is closely related to and, in some cases, can be explained by satisficing theory (Krosnick, 1991) and possibly, social desirability bias. Participants may carefully consider the numerous explanations for choosing one option over the other or simply respond to the item in a way that is most socially desirable (i.e., the chosen option displays the respondent in a positive and admirable way). Acquiescence is most common when participants do not enjoy high levels of thinking or exerting a high amount of cognitive workload. These tendencies may be due to question difficulty, disinterest in the survey topic (Groves, Presser, & Dipko, 2004), or simply because participants are inconvenienced by responding to the questionnaire (Krosnick, 1999).

Krosnick (1999) provided an additional possible issue concerning the amount, or level, of cognitive processing involved in responding to items on a questionnaire. Explicitly, participants progress through a series of steps during the survey process in order to optimize their response. First, participants must read the question, interpret its meaning, and infer its intent. Next, participants must assess its relevancy to their own personhood and assimilate that information into their current state, environment, or situation. Last, participants must render a response to the questionnaire item based on the aforementioned process of information. The final judgment is a result of the amount of cognitive workload expended (Heerwegh & Loosveldt, 2008; Karabenick et al., 2007). Ideally, participants would expend a great deal of cognitive workload, optimizing their responses to questionnaire items. However, issues arise if items are too difficult to understand or interpret, require a great deal of cognitive effort, perceived as too personal, are simply uninteresting, or create a burden on the participant (Krosnick, 1999).

Heerwegh and Loosveldt (2008) and Karabenick et al. (2007) determined how participants respond to items on a questionnaire depends on the amount of cognitive workload that is required by individual survey items. This assumption is a basic tenet of ‘satisficing theory’ (Krosnick, 1991). While some participants will expend considerable effort in responding to items, others will use “cognitive shortcuts” to reduce the amount of required effort thus demonstrating

satisficing behavior, especially in attitudinal measures (Krosnick, 1991). How participants respond is based on their motivation and ability to respond to the difficulty level of the survey items (or tasks). If participants are highly motivated and able to respond with little difficulty, satisficing will be low. As the difficulty of the task (or item content) increases, satisficing behavior increases. That is, the more difficult the task, the higher the likelihood respondents will use cognitive shortcuts thereby potentially reducing the quality of data. That is, the way a participant responds may be due to the lack of effort spent on the task (Krosnick, 1991).

Heerwegh and Loosveldt (2008) noted the administration mode may influence a participant's likelihood to exhibit satisficing behavior. Multiple distractions (*described earlier*) may prevent participants from fully engaging in the survey completion process. This is particularly true for web-based surveys. Some questionnaires require a considerable amount of cognitive effort or demand which increases satisficing behavior. Heerwegh and Loosveldt (2008) found when given a response choice of "don't know," participants who responded to an online survey did so at a higher rate than participants who responded in a paper-and-pencil format. Joinson, Woodley, and Reips (2007) suggested a response option of "I don't know", "I prefer not to answer," or "No opinion" may improve data quality by allowing respondents to complete a survey by default. That is, allowing participants to respond to all items on a questionnaire without disclosing personal or sensitive information results in a higher response rate and overall improves data quality.

Djamasbi, Siegel, and Tullis (2010) found how surveys are designed (i.e., aesthetic appeal) influences how participants respond to items on a questionnaire. Specifically, an aesthetically pleasing design can foster a participant's attention to complete the survey, resulting in a higher response rate. Schmidt, Liu, and Sridharan (2009) explored the effects of aesthetics on webpage usability and found users preferred visually appealing web pages while Shropshire, Hawdon, and Witte (2009) supported the use of visual enhancements in order to increase response rate. Taken together, the various aspects of how surveys are constructed and delivered

influence how participants respond to requests for participation in survey research as well as how participants respond to items on online questionnaires.

Computer self-efficacy

Compeau and Higgins (1995) contributed to the development of a theory toward computer self-efficacy, broadly defined as "a judgment of one's capability to use a computer." More specifically, the authors were interested in participants' reactions to the many facets of computer technology. These reactions, according to Compeau and Higgins, are based on participants' perceived ability to use a computer based on previous experience and confidence in future accomplishments. Conrad and Munro (2008) stressed the rapid, continuous shifts in both computer technology and the Internet requires researchers to maintain a current measurement of computer self-efficacy. However, Weigold, Weigold, and Russell (2013) maintained the appropriateness of the Computer Self-Efficacy Scale (CSE)—developed by Murphy, Coover, and Owen (1989)—due to its popularity, or use, in multiple populations. Several other measures of computer self-efficacy are available and have produced favorable reliability and validity (Compeau & Higgins, 1995; Conrad & Munro, 2008; Heinssen, Glass, & Knight, 1987; Kay, 1989; Gressard & Loyd, 1985; Meier, 1988) but were developed prior to the emergence of the Internet (Torkzadeh, Chang, & Demirhan, 2006) and conceptualized to measure different dimensions of computer technology (Conrad & Munro, 2008). Researchers should carefully consider which domains of computer self-efficacy they are interested in studying and choose the appropriate measurement accordingly.

The increasing reliance on technology and web-based communication (e.g., the Internet) to interact with others has all but forced Internet users to become more proficient in the various capabilities of computers and computer-based programs (Jones, Yale, Millermaier, & Pérez, 2008). Perhaps the most prevalent demographic group disposed to using technology—personal computers, tablets and e-readers, and mobile devices— for the majority of their professional and personal interaction are college students. Universities and colleges are increasingly relying on

computers and their capabilities (Kuenzi, 1999-2000). Correspondingly, college students need to be well-adept so as to facilitate their ability to understand the software programs and technology prevalent in the college curriculum (Messineo & DeOllos, 2005; Torkzadeh & van Dyke, 2002). Weston and Barker (2002) examined student computer use and levels of proficiency and found students rated themselves as adequate to highly adept in terms of their ability to use various operating systems and hardware and software programs as well as use the Internet for email, online coursework, and scholarly research.

Several investigations have examined the perceptions of higher education students regarding their level of computer use or experience (Jones, Yale, Millermaier, & Pérez, 2002; Weston & Barker, 2002), comfort (Kinzie, Delcourt, & Powers, 1994), skill (McCoy, 2010; Messineo & DeOllos, 2005**;** Weston & Barker, 2002), and perceived computer self-efficacy (Kinzie, Delcourt, & Powers, 1994; McCoy, 2010). Previous research has shown that while college students may rely heavily on the Internet—namely, email— to interact with their professors and partake in online courses (Jones, Yale, Millermaier, & Pérez, 2002) and also report high levels of experience with Internet, email, and word-processing programs, they face some difficulty when encountering more advanced technology (Messineo & DeOllos, 2005). According to Kinzie, Delcourt, and Powers (1994) and McCoy (2010), less adept users may lack the confidence due to less experience resulting in a negative attitude towards computer technology which in turn influences their overall computer self-efficacy. Torkzadeh, Chang, and Demirhan (2006) recommended computer and Internet training, specifically task-specific exercises, may improve one's self-efficacy and promote a more positive attitude toward technology.

Students who lack the necessary technical knowledge, skills, and abilities to perform the requirements of many college courses may encounter some anxiety if exposed to unfamiliar or novel information technology systems (Messineo & DeOllos, 2005). As a result, a student's self-efficacy and level of expected academic achievement and performance may be diminished. Self-efficacy theory was conceptualized by Bandura and Wood in 1989 as a person's perception of his

or her capabilities regarding a certain task depending on the individual's current level of motivation, cognitive abilities, and the elements of the task. Previous research has demonstrated that individuals with low self-efficacy tend to disengage or abstain from tasks that are too difficult or are believed to be unachievable (Bandura, 1978; Bandura, 1982). Conversely, high self-efficacious people tend to endure challenging activities, despite the task's difficulty (Torkzadeh & van Dyke, 2002). Students who perceive their ability is adept enough to meet the increasing "situational demands", especially those associated with Internet use (Moos & Azevedo, 2009), are influenced by their previous performance on similar tasks, how motivated they are to persist in the face of difficulty, and their willingness to engage in such activity.

Oliver (2002) outlined four primary issues in user perception of online web pages: usability, visualization, functionality, and accessibility. Usability—how webpage information is navigated and processed—is related to the overall aesthetics, or visualization, of the web page which is simultaneously and intricately linked to both accessibility—visual design tools such as font type, font size, and font color—and functionality, or the appropriateness of design features. Schmidt, Liu, and Sridharan (2009) found webpage aesthetics, or the design and layout of a webpage, influences viewers' (i.e., web users') ability to process the information on the web page. That is, a higher amount of cognitive processing is required by the user for web pages with highly sophisticated or visually enhanced designs. Djamasbi, Siegel, and Tullis (2010) investigated the importance of enhanced visualization, or appeal, of the web and found younger, more adept Internet users—specifically, Generation Y members—prefer, and often expect, aesthetically pleasing web pages. The authors attribute this tendency to this age group's experience with the Internet and its capabilities. That is, Generation Y members have been exposed to technological advances from an early age and are thus generally considered to be highly adept at perusing various web pages (Messineo & DeOllos, 2005).

Summary

Previous research has shown inconsistent findings on the equivalence of paper-based and online modes of survey administration (Webster & Compeau, 1996; Weigold, Weigold, & Russell, 2013). While some studies have shown equivalence based on psychometric properties (Brock, Barry, Lawrence, Dey, & Rolffs, 2010; de Beuckelaer & Lievens, 2008; Gwaltney, Shields, & Shiffman, 2008; Perkins, 2004; Ryan, Corry, Attewell, & Smithson, 2002; Sethuraman, Kerin, & Cron, 2005; Weigold, Weigold, and Russell, 2013), other analyses (Couper, Traugott, & Lamias, 2001; Webster & Compeau,1996) have indicated equivalence is influenced by how questionnaire items are presented to respondents. Several investigations (Couper, Traugott, & Lamias, 2001; Redline & Dillman, 1999; Ware, 2000) have found that while both paper-and-pencil and Internet surveys are self-administered and rely on both verbal (i.e., order of questionnaire items and question wording) and visual (i.e., design elements) information to interact with participants, the responses of participants are inevitably influenced by the way the survey is organized, both verbally and visually (Couper, Traugott, & Lamias, 2001).

Deutskens, Ruyter, Wetzels, and Oosterveld (2004) found visually enhanced designs may improve response quality. Extending previous research on the impact of visualization on task completion, Garcia and Caldera (1996) explored the effect of color and typeface on the readability of online text and found contrasting colors—light background color with dark colored foreground text—influenced readability (as measured by reading speed). Therefore, researchers may consider using less conventional color choices in the aesthetic design of an online questionnaire. Garcia and Caldera (1996) focused their efforts on the interaction of text and background and how information appears visually to respondents. The authors reiterated the need for studying the psychological implications for using a variety of colors when presenting information, especially in an online format.

An extensive search of the literature reviewing the effects of visual design of online surveys revealed that not much is known about the comparability of different online survey

design formats. That is, there is little to no research concerning the impact of different online survey modes (or design formats) of administration. Variation in the design formats of online or web-based questionnaires has generally indicated participants' responses are influenced by visual effects (Djamasbi, Siegel, & Tullis, 2010). New software tools have made it possible for researchers to vary the way in which questionnaires are designed and presented in an online format. Wright (2005) provided a list of several software packages available to researchers for the online collection of data. These services provide effective tools in both the contextual and visual design of the online survey. These design options can range from basic standard features to a highly sophisticated design organization (Couper, Traugott, & Lamias, 2001).

Current study

The aim of this current study is to assess the effects of survey design on participants' responses to online questionnaires, presented in different yet equivalent formats (Lietz, 2010) or modes (Couper, 2011). Best and Krueger (2004) found most surveys include a design using one of two widely accepted survey typefaces—Serif (e.g., Times New Roman and Courier New) or San Serif (e.g., Arial)—and present black text on a white background. This format is typically used due to its high legibility (Garcia & Caldera, 1996). Sue and Ritter (2007) found participants are able to read online script faster when verbal information is presented in this format. Furthermore, Redline, Dillman, Carley-Baxter, and Creecy (2003) suggested visualization of online surveys should be relatively simple, sophisticated and visually appealing in order to engage participants to remain focused on the information (i.e., text) on the screen. Such design elements are important in understanding the how participants respond to questionnaire items. Too many design elements may result in cognitive overload and, subsequently, non-completion, while very basic designs (e.g., black text on white background) may lead to boredom or disinterest, also resulting in a low response rate.

Hartley and Rutherford (2003) assessed the effects of using colored paper—in paper-based questionnaires and survey design—on response rates and found superior rates of

responding for studies using light colored paper such as pink. Darker colors, such as green and blue, were believed to produce negative attitudes toward the questionnaire and did not produce higher response rates. In an examination of the effects of color—background and text—on participants' reactions to unconventional color schemes, Godar (2000) found some color combinations elicited lower responses to items on a questionnaire. Furthermore, Hartley and Rutherford (2003) found previous research had not considered the various aspects of color—saturation, brightness, and hue—available in online questionnaire design (Couper, Traugott, & Lamias, 2001). Taken together, the effects of using colored paper for paper-based administration may not directly translate into other contexts, or other modes of survey administration such as web-based questionnaires (Godar, 2000).

Couper (2011) advocated for the use of multiple modes to collect data providing the modes are designed similarly, or with few variations in how the questionnaire is presented. As survey modes evolve and integrate novel ways of designing and administering questionnaires, research investigations into the effects of survey mode will be expected to develop new techniques and strategies to meet these societal changes. Regardless of the administration mode, questionnaires must be designed in a professional, aesthetically-pleasing way (Walston, Lissitz, & Rudner, 2006). Using this technique, participants will perceive not only a serious, business-like tone to the research topic but encounter a visually appealing questionnaire. Experienced technology users—of computers and the Internet—may be more capable of discerning these visual design elements. Respondents to more sophisticatedly-designed online surveys may also appreciate such visual effects more than an inexperienced, low efficacious user. That is, a participant whose computer self-efficacy is low may experience a level of anxiety when encountering an intricately designed questionnaire.

CHAPTER III

METHODOLOGY

This chapter describes the methods that were used to complete this research. The participants are described, as well as the procedures used for recruitment. The design of the study, including the study's independent and dependent variables, is also described. Also included in this section is an introduction to the instruments that were used for data collection. Procedures for data collection and data analysis are also reviewed.

Participants

The 216 participants in this study were Oklahoma State University (OSU) students, enrolled at the OSU-Stillwater campus during the Fall 2013 semester. Established in 1890, OSU now enrolls more than 35,000 students across its five-campus system (OSU-Stillwater, OSU-Oklahoma City, OSU-Tulsa, OSU Institute of Technology, and OSU Center for Health Sciences in Tulsa) with approximately 23,033 students enrolled at the OSU-Stillwater campus in the Fall of 2013 (OSU Institutional Research and Information Management [IRIM], 2013).

As a modern land-grant institution, OSU is committed to "improve[ing] the lives of people in Oklahoma, the nation, and the world through integrated, high-quality teaching, research, and outreach" (OSU System, 2013). Its mission "to advance knowledge, enrich lives, and stimulate economic development through instruction, research, outreach, and creative activities" is evidenced by the university's strong undergraduate and graduate programs.

A request was made to the OSU IRIM office for a list of student email information. The maximum number of student records IRIM is allowed to distribute, as per the guidelines of the OSU Office of Communications and the OSU Institutional Review Board (IRB), is 5,000. For this study, a list of 4,800 student emails was requested so as to evenly distribute the sample among the 16 testing conditions. This list consisted of a random sample of both undergraduate and graduate students enrolled at OSU-Stillwater. No other exclusion criteria were used for this study.

The list of emails provided by IRIM was delivered to the researcher's school email address in an Excel spreadsheet. Prior to distributing the survey, the list of 4,800 email addresses was divided into 16 groups of 300 emails each. The groups were developed using a systematic process in which every nth email address was assigned to one of the 16 groups. More precisely, the first name on the list was assigned to condition 1; the second name on the list was assigned to condition 2, and so on. This method of assignment cycled until all participants were assigned to a testing condition (the 17th name on the list was assigned to condition 1; the 18th name on the list was assigned to condition 2, and so on). After the list of 4,800 emails had been catalogued into 16 groups of 300 emails each, the researcher's personal Microsoft Word software was used to distribute the emails using the program's Mail Merge feature (described below). In order to use the Mail Merge feature, the researcher had to sync her personal Microsoft Word 2010 software with the Microsoft Outlook/Exchange 2010 program on her computer. (Microsoft Outlook/Exchange serves as an email provider and is offered to all OSU students. To facilitate a more efficient method of sending recruitment emails to participants, the researcher set up an account using her OSU student email address. This process allowed the email with the invitation to participate in the study to be sent from the researcher's student email address via Microsoft Word).

Sixteen surveys, each of which varied in visual design characteristics but contained the same content, were created in Qualtrics. Upon creating each survey, a link was generated for that

survey. Thus, 16 links were produced, one for each of the 16 surveys. These individual links were then copy and pasted into solicitation emails specific to the testing condition in which the participant had been assigned. That is, for each group, the solicitation email contained a general description of the study as well as the unique link for that group.

As stated before, participants had been assigned to a specific testing condition using a systematic process in Excel. This process produced a recipient list to be used during the Mail Merge process in Microsoft Word. (The Mail Merge feature allows email messages, letters, and other documents to be sent to many recipients. It is able to utilize a list, typically created in Excel workbook, in order to send messages to multiple recipients at the same time.). In order to ensure all of the 16 groups received the survey link specific to that group, the Step-by-Step Mail Merge procedure was used to send each of the 16 solicitation emails. Therefore, 16 solicitation emails were sent in groups of 300 by means of the Step-by-Step Mail Merge Wizard feature in Microsoft Word. Thus, using this method, all students were contacted via email at their OSU email address with the link specific to their assigned group membership inviting them to participate in the study.

Design

This study consisted of a 2 x 2 x 2 x 2 completely randomized factorial ANOVA design. A description of the levels is described as follows: 1) color of font: red or blue, 2) color of background: white or black, 3) progress bar indicator: present or absent, 4) highlight of item currently addressing: present or absent. There were 16 different combinations in which these variables were manipulated according to their corresponding levels. The combinations were divided into four different sets in which the color of font and the color of background were manipulated by varying the color of the font according to the corresponding background. Both progress bar and highlight bar were either present or absent depending on their corresponding condition.

Independent variables

Four independent variables were used in this study. The independent variables were 1) text (or font) color, 2) background color, 3) use of a highlight bar, and 4) use of a progress bar indicator. Each of the independent variables contained two (2) levels. These variables were selected based on their relatively prevalent use in online survey design. Specifically color, both font and background, is a feature of online survey design that is readily apparent and easily manipulated. Furthermore, highlight bars and progress bar indicators are features unique to online survey design. The absence or presence of each of these variables serves to identify the merit, or value, of these features.

Text color. The text color variable was manipulated by varying the color of font presented to the participants. The two colors selected for this study were red and blue. These colors were chosen based on their legibility (Lynch & Horton, 2009) as well as their degree of readability and level of contrast to a black or white background (Kyrnin, n.d.).

Background color. The background color was manipulated by varying the color of background. The two colors selected for the background condition were black or white. These colors were selected based on their high level of contrast to the text colors chosen for this study. According to Cannon (2012), using contrasting colors can help guide the viewer's attention to certain elements on the computer screen.

Highlighting items. Questions were either highlighted or not highlighted, depending on the condition in which the participant was assigned. The highlight bar was used as an indicator to allow the participant to view the item to which they are currently responding. The color of the highlight bar was yellow.

Progress bar. The progress bar indicator was either present or absent from the survey, depending on which condition the participant was assigned. The progress bar was located at the end of each page. Participants were able to view their current completion of the survey on a '0' to

‘100’ percentage rate. Although participants were not shown an exact percent of their progress, they were able to estimate their completion using the bar.

Concisely, a total of 16 variable combinations comprised the testing conditions. The conditions are presented below:

Table 1.

Survey Conditions

				Background color			
				Black		White	
				Progress bar			
				Yes	No	Yes	No
Color of font	Red	Highlight bar	Yes	1	5	9	13
			No	2	6	10	14
	Blue		Yes	3	7	11	15
			No	4	8	12	16

Dependent variables

There were five dependent variables included in this study: 1) response rate, 2) completion time, and differences among group conditions on 3) a government knowledge test, 4) an attitudinal measure, and 5) a computer self-efficacy scale. Response rate was measured by number of participants who responded to and completed the survey. Survey completion time included the amount of time it took participants to complete the survey. Differences among groups included scores on all three instruments—knowledge test, attitudinal questionnaire, and computer self-efficacy scale—used in the research study. Regardless of the visual presentation, all participants viewed the same questionnaire items (i.e., content).

Instruments

Three instruments were used in this study: 1) a knowledge test, 2) an attitudinal measure, and 3) a computer self-efficacy questionnaire. The effects on participants' reported self-efficacy were evaluated through scores on the knowledge and attitude measures.

Knowledge test

A knowledge test was used for two purposes: 1) to assess participants' attention to the survey, and 2) to further understand the effects of color on the legibility and readability of online text. The knowledge test consisted of 20 items related to American History and Government (e.g., Civics; See Appendix A). These questions were randomly sampled from the Civics portion of the United States Citizen and Immigration Services' test for American Citizen Naturalization[1]. In order to determine the 20 items that were to be used on the knowledge test, a pilot study was conducted.

It was assumed participants had a basic knowledge of United States history and government. If participants are attending to the test and can legibly read the knowledge items, 100% accuracy is expected on this portion of the survey process. Therefore, this measure served as a basis for evaluating whether participants were fully engaged in the research process by asking general education questions. The pilot test also served as gauge against which questions may not be answered correctly.

Pilot study for the Government knowledge test

Prior to the onset of the current study, a small pilot study was conducted in order to determine which Civics questions would be most appropriate for use on the final form. Thirty questions were selected for the pilot study survey based on the level of knowledge necessary to answer the questions correctly. Questions that were answered correctly by 26 out of the 30 pilot

[1] The actual Civics test is presented in an interview format in which a USCIS Officer asks prospective naturalized citizens up to 10 questions from a list of 100 questions. Test-takers must verbally respond to and correctly answer six questions in order to pass the Civics portion of the Citizen Naturalization test.

study respondents were retained for use on the final test. The final Knowledge test consisted of 20 questions on the subject of American History and Government.

During the Summer 2013 school session, a small sample of students was recruited from a General Psychology course at Southwestern Oklahoma State University to participate the pilot study. Students were invited to participate via an email providing general information about the study. Students who desired to participate in the study were instructed to contact the researcher at her student email address for a link to the survey. (The 30-item survey had been created in Qualtrics. A link, specific to the survey, was generated for distribution to students who requested the link to participate). Upon completion of the study, students were instructed to send an email to the researcher with their name in the body of the email message in order to receive course credit for participating.

Attitudinal measure

An attitudinal measure served as a basis to evaluate differences in scores across testing conditions. The instrument chosen for this portion of the survey included items assessing undergraduate level students' attitudes toward research methods courses. Items comprising the scale included usefulness of research in participants' careers, anxiety and stress of understanding research, positive attitude towards research, relevance of research to participants' professional and personal lives, and having trouble with research (See Appendix B).

The "Attitudes Toward Research" (ATR) scale was developed in 2002 by Papanastasiou as a means to help instructors ease the anxiety students often experience when learning the research process and create a more positive experience with and attitude toward research (Papanastasiou, 2005). The initial ATR consisted of 56 attitudinal items, both positively- and negatively- worded, in which participants responded on a 7-point Likert scale. Negatively-worded items were reverse scored so that high number responses would reflect positive attitudes. Further analyses revealed 24 "inappropriate" items in the original scale. Factor analysis of the 32-item ATR exposed a five factor solution accounting for 66.25% of the total variance.

Responses on the ATR in 2002 indicated a high reliability for the 32 items on the test (Cronbach's alpha = 0.948). Coefficient alphas were calculated for the responses to items on each of the five subscales: research usefulness in the profession factor was 0.919 (9 items); research anxiety was 0.918 (8 items); positive attitude towards research was 0.929 (8 items); relevancy of research toward personal life was 0.767 (4 items); and research difficulty was 0.7111 (3 items).

The interrelationships of the five subscales of the ATR suggested the 'usefulness' factor is highly correlated with 'relevancy towards life' factor (r = 0.69) and with 'positive attitudes toward research' factor (r = 0.67). The 'anxiety toward research' scale was most highly correlated with 'positive attitudes toward research' (r = .58) and 'research difficulty' (r = .52) scales. 'Research difficulty' was also correlated with 'research anxiety' (r = .52).

Computer self-efficacy

The computer self-efficacy measure consisted of an instrument commonly used to evaluate participants' self-reported use, experience with, attitudes concerning, and anxiety toward the many different facets of computers, computer programs, and the Internet (see Appendix C).

The Computer User Self-Efficacy scale (CUSE; Cassidy & Eachus, 2002) is comprised of two parts: 1) a short demographic survey in which participants are asked to provide basic background information including their experience with computers, and 2) a 30-item questionnaire in which participants were asked to indicate their strength of agreement with statements about their attitudes toward computers. Participants responded to the items on a 6-point Likert scale.

Findings from a development and validation study by Cassidy and Eachus (2002) indicated high internal consistency of the CUSE, as measured by Cronbach's alpha (α= 0.97, N = 184).Test-retest reliability over a one-month period was statistically significant (r = 0.86, N = 74, p < .0005.). External validity was reported at a satisfactory level (r = 0.86). The CUSE has been related to computer experience, familiarity with software packages, computer training, and computer ownership. Correlations between self-efficacy and experience were statistically

significant (r = 079, p < .0005, N = 212). Familiarity was also highly correlated with self-efficacy (r = 0.75, p < .0005, N = 210).

The CUSE has chiefly been used by researchers to assess the computer self-efficacy in the adult population, especially for college and university students. The reliance of technology in both the academic and administrative sectors of educational institutions has led researchers to assess the impact of industrial changes on computer users' ability to adapt to these changes.

Demographic Data

The final phase of data collection included a request for participants to respond to a short demographic inventory (See Appendix D). Questions appearing on this list of items included participant's college classification and college major. Participants were also asked to estimate the amount of time, including time spent on the Internet, they spend daily on the computer for either professional or educational purposes. Two items requested participants to specify the conditions in which they are viewing the survey: 1) type of device (e.g., Mac, PC, tablet, mobile phone, etc.) used to complete the survey, and 2) currently wearing contacts or glasses. Participants were asked to provide their thoughts about the survey process, including the visual design elements of the survey.

Procedures

This study was conducted using Qualtrics, an online survey software program for data collection and analyses. Each survey had a link specific to that survey. Solicitation emails were sent from Microsoft Outlook via the Step-by-Step Mail Merge Wizard in Microsoft Word Participants were emailed the solicitation email which contained general information about the study as well as a link specific to the corresponding testing condition. That is, the link directed the participants to their customized survey. Initial requests for participation were made in the 3rd week of the Fall 2013 semester. Because Oklahoma State University limits the number of reminder emails sent to participants to one, a second email was sent to all participants three weeks after the initial invitation to participate. This reminder email allowed those who had not yet

responded to the request to participate in the study. Participation in the study was completely voluntary. Students who chose to participate in the study were allowed to cease participation at any time without penalty.

CHAPTER IV

RESULTS

The purpose of this study was to extend the literature on the effects of variations in the visual representation of online survey designs. Specifically, it aimed to assess the effects of online survey design characteristics on participants' responses to online questionnaires presented in different formats. This chapter is a presentation of the results of the analysis of data collected from 216 participants in a study of the effects of color (i.e., background color and font color) and visual aids (i.e., progress bar and highlight bar) on how participants respond to items on a questionnaire. The analyses were guided by two research questions:

1. Does text color, background color, appearance of progress bar, or appearance of highlight bar affect the a) response rate, b) completion time, and group scores on c) a knowledge test, d) attitudinal measure, and e) self-efficacy scale when administered in a web-based mode?
2. Are there significant interactions of text color, background color, appearance of progress bar, or appearance of highlight bar on a) response rate, b) completion time, and group scores on c) a knowledge test, d) attitudinal measure, and e) self-efficacy scale when administered in a web-based mode?

The following pages are divided into the following sections. The first section will describe the characteristics of the sample, including demographic information collected in part one of the Computer User Self-Efficacy scale. The second section will describe the design of the study, including the quantitative statistical analyses used to answer each research question as well as the findings for each research question and its analysis. The third section will describe the qualitative analyses performed on the textual data—feedback obtained from participants about the visual characteristics of their assigned survey. The chapter concludes with a summary of the findings.

Sample characteristics

Of the 4,800 student emails provided by OSU IRIM, a total of 290 responses were recorded by Qualtrics. Preliminary analysis of the data began by extracting the data from Qualtrics, the computer software program from which the data were collected, and verifying that the degree of completion for each survey was more than one-third of the total survey (i.e., participants responded to at least 32 of the 93 survey items). From the initial 290 responses, 74 responses were excluded from the total dataset based on the following criteria: students who either clicked on the link to the survey site but did not click past the IRB consent screen or "dropped out" of the study after completing less than one-third of the survey (i.e., data obtained from participants who dropped out after completing the 'Attitudes toward Research' measure only were not included in the analyses). The total number of responses after excluding the missing data was 216. With a total of 216 participants, the distribution of participants among the 16 groups was unequal. Table 2 shows the frequency of responses to each of the 16 testing conditions as well as the response rate of each survey, both prior to and after excluding the missing data.

Table 2.

Comparison of response rates

Testing condition	# completed surveys*	# total responses	Percentage complete
1	7	13	53.8%
2	9	11	81.8%
3	11	18	61.1%
4	14	20	70.0%
5	11	14	78.6%
6	21	27	77.8%
7	16	20	80.0%
8	18	22	81.8%
9	14	17	82.4%
10	10	12	83.3%
11	13	19	68.4%
12	16	20	80.0%
13	7	13	53.8%
14	13	16	81.3%
15	19	22	86.4%
16	17	26	65.4%

*Note: # *of completed surveys* refers to the number of surveys that were determined to be more than two-thirds, or 67 percent, complete.

Participants in the sample included both undergraduate and graduate students enrolled in the fall 2013 semester at Oklahoma State University (see Table 3). The age of the participants ranged from 18 to over 53, with 74.5% of students (n=161) between the ages of 18 and 24. Thirty-five participants (16.2%) were aged between 25 and 31 years. The remaining 9.2% participants were classified as follows: nine participants (4.2%) between the ages of 32 and 38;

five participants (2.3%) between the ages of 39 and 45; two participants (0.9%) between the ages of 46 and 52; and four participants (1.9%) aged 53 and older. The population consisted of 60 freshman (27.8%), 35 sophomores (16.2%), 26 juniors (12.0%), 35 seniors (16.2%), and 54 (25.0%) graduate students.

Table 3.

Sample characteristics and demographics

Age		
	18-24	161 (74.5%)
	25-31	35 (16.2%)
	32-38	9 (4.2%)
	39-45	5 (2.3%)
	46-52	2 (0.9%)
	53+	4 (1.9%)
		n = 216
College classification		
	Freshman	60 (27.8%)
	Sophomore	35 (16.2%)
	Junior	26 (12.0%)
	Senior	35 (16.2%)
	Graduate	54 (25.0%)
		*n = 210**
Time spent on computer		
	Less than one (1) hour	2 (0.9%)
	1 to 3 hours	79 (36.6%)
	4 to 6 hours	74 (34.3%)

	7 to 9 hours	39 (18.1%)
	9 or more hours	16 (7.4%)
		*n = 210**
Device used to complete survey		
	Personal computer (desktop)	44 (20.4%)
	Personal computer (laptop)	134 (62.0%)
	Tablet	1 (0.5%)
	Mobile phone	22 (10.2%)
	Other	8 (3.7%)
		*n = 209**
Are you currently wearing glasses or contacts?		
	Yes	105 (48.6%)
	No	104 (48.1%)
		*n = 209**

*Note: Missing data. Demographic information was collected in the final portion of the survey. Providing this information was optional; therefore, participants who chose not to reveal this information would not be included in this table. Demographic data from participants who "dropped out" prior to finishing the survey would also not be included here.

When participants were asked about how much time they spend on the computer for either professional/educational or personal use, the majority of participants (n=79) reported they spent one to three hours on the computer while 74 participants reported they spent four to six hours on the computer each day(see Table 2). One hundred and thirty four participants (62%) completed the survey on a personal laptop computer while 20.4% of participants (N=44) completed the survey on a personal desktop computer (see Table 3). Participants were also asked if they were currently wearing glasses or contacts (i.e., wearing glasses or contacts while taking the survey). Of the 209 (96.8%) recorded responses, 105 (48.6%) reported they were wearing glasses or contacts while 104 (48.1%) reported they were not wearing glasses or contacts.

Table 4

Responses to Part 1 of the Computer User Self-Efficacy (CUSE) scale

Experience with computers		
	None	1 (0.5%)
	Limited	3 (1.4%)
	Some	66 (30.6%)
	Quite a lot	103 (47.7%)
	Extensive	43 (19.9%)
		n = 216
Computer packages (software) used*		
	Word processing	206 (95.4%)
	Spreadsheets	176 (81.5%)
	Databases	126 (58.3%)
	Presentation packages	147 (68.1%)
	Statistics packages	59 (27.3%)
	Desktop publishing	80 (37.0%)
	Multimedia	139 (64.4%)
	Other (please specify)	21 (9.7%)
Do you own a computer?		
	Yes	213 (98.6%)
	No	3 (1.4%)
		n = 216
Have you ever attended a computer training course?		
	Yes	138 (63.9%)

	No	77 (35.6%)
		*n = 215***
Age****		

*Note: Participants were asked to indicate the types of computer (software) packages they have used. Frequencies and percentages reflect those participants who responded to this question; not all participants replied. Some participants may have selected more than one option (i.e., computer package).
**Note: One participant did not report whether they had attended a computer training course.
***Note: Data collected for 'Age' is presented in *Table XX. Sample characteristics and demographics*.

Part One of the Computer User Self-Efficacy (CUSE) scale asked participants to provide some basic information about their experiences, if any, of computers. Of the 216 responding participants, 213 indicated that they owned a computer; 138 reported that they had attended a computer training course; and 103 revealed "quite a lot" experience with computers. Participants were also asked to indicate the computer packages or software they have used. See Table 4 shows the responses of participants for Part One of CUSE scale.

Design of the study

The analysis of data was designed to determine whether or not variations in the visual presentation of an online survey influenced how participants responded to questions. In each of the 16 different conditions, participants observed a combination of the following design aspects of online survey development: a) font color, b) background color, c) highlight bar, and d) progress bar indicator. Table 5 represents the 16 conditions. The conditions were designed with the purpose of revealing statistically significant differences between survey conditions. That is, whether or not variations in the presentation of online survey design influenced how participants responded to questions in an online survey. Participants were assigned to one of 16 testing conditions in which the visual presentation of the survey differed in terms of font color (i.e., red or blue), background color (i.e., black or white), highlight bar (i.e., present or absent), and progress bar (i.e., present or absent). Below is a table demonstrating the testing conditions:

Table 5.

The 2x2x2x2 Completely Randomized Factorial (CRF) ANOVAs

				Background color			
				Black		White	
				Progress bar			
				Yes	No	Yes	No
Color of font	Red	Highlight bar	Yes	1	5	9	13
			No	2	6	10	14
	Blue		Yes	3	7	11	15
			No	4	8	12	16

Dependent variables

This study included five dependent variables. Due to the relative independence of the three instruments (Government knowledge test, Attitudes toward Research [ATR] scale, and Computer User Self-Efficacy [CUSE] scale) used in the study, a series of four 2x2x2x2 completely randomized factorial (CRF) analysis of variances (ANOVAs) were performed in order to determine the effects of the independent variables—font color, background color, highlight bar, and progress bar indicator—on the dependent measures. More precisely, a 2x2x2x2 CRF-ANOVA was performed for each of the three instruments to determine if there were statistically significant differences in responses among the 16 groups, or testing conditions. A 2x2x2x2 CRF-ANOVA was also performed for completion time—as measured by the mean completion time for each group—to determine if there were statistically significant differences among groups for the total amount of time it took to complete the survey.

Analyses of the data sought to identify the differences among the 16 conditions and were guided by the question of whether or not there was a significant interaction for a) font color, b) background color, c) highlight bar, or d) progress bar on scores on each of the three instruments

as well as for completion times of the 16 groups. (Analysis of the data under the response rate condition is discussed more fully in the 'Qualitative Analyses' section). Table 6 presents a detailed summary of the main effects, two-way interactions, three-way interactions, and four-way interactions of the four 2x2x2x2 ANOVAs conducted in the study.

Table 6.

Summary of ANOVAs for dependent measures

	Dependent measures											
	Knowledge test (Government)			ATR			CUSE			Completion time		
	df	F	*p*	*df*	F	*p*	*df*	F	*p*	*df*	F	*p*
Main effects												
Font	1	3.192	.075	1	.099	.753	1	.581	.447	1	2.961	.087
Progress bar (P_Bar)	1	.397	.529	1	.036	.850	1	4.764	.030	1	1.416	.235
Highlight bar (H_Bar)	1	.580.	.447	1	1.466	.227	1	3.256	.073	1	3.563	.061
Background color (Back_Color)	1	.205	.651	1	3.948	.048*	1	.311	.578	1	1.607	.207
Two-way interactions												
Font x P_Bar	1	.018	.893	1	.097	.755	1	.223	.637	1	3.811	.052
Font x H_Bar	1	.255	.614	1	1.176	.280	1	.940	.333	1	.120	.729
Font x Back_Color	1	.168	.683	1	.043	.836	1	6.844	.010*	1	.943	.333
P_Bar x H_Bar	1	.832	.363	1	2.347	.127	1	.110	.741	1	.000	.989
P_Bar x	1	2.447	.119	1	.004	.948	1	.608	.436	1	1.206	.274

Back_Color												
H_Bar x Back_Color	1	2.134	.146	1	.101	.751	1	2.828	.094	1	.135	.714
Three-way interactions												
Font x P_Bar x H_Bar	1	2.753	.099	1	.384	.536	1	.016	.900	1	.210	.648
Font x P_Bar x Back_Color	1	.670	.414	1	.009	.924	1	.098	.754	1	1.337	.249
Font x H_Bar x Back_Color	1	.000	.985	1	.015	.902	1	1.026	.312	1	1.284	.259
P_Bar x H_Bar x Back_Color	1	1.417	.235	1	.066	.797	1	.267	.606	1	1.045	.308
Four way interaction												
Font x P_bar x H_bar x Back_color	.069	1	.793	3.731	1	.055	1.535	1	.217	3.594	1	.060

*Note: Significant at the $p<.05$ level.

Analysis of the data under the mean differences among groups condition was guided by the question of whether or not there was a significant interaction for a) font color, b) background color, c) highlight bar, and d) progress bar among groups based on participants' responses to the each of the three instruments—Government knowledge test, 'Attitudes toward Research,' Government test, and 'Computer User Self-Efficacy'—used in the study. The following sections will discuss the individual results of these analyses.

Knowledge instrument: Government test

A 2x2x2x2 CRF-ANOVA was performed in order to determine whether there was a statistically significant interaction among font color, background color, highlight bar, and progress bar indicator on a Knowledge test containing American History and Government items. Results from analyses of the Knowledge test revealed no significant main effects or interactions among font color, background color, use of a highlight bar, and use of a progress bar indicator on responses to items on the Knowledge test.

Attitudes toward Research

A 2x2x2x2 CRF-ANOVA was performed in order to determine whether there was a statistically significant interaction among font color, background color, highlight bar, and progress bar indicator on responses to the Attitudes toward Research scale. Statistical analyses also showed a significant main effect (p=.048) for background color [F(1,211) = 3.948]. The significant main effect was associated with a partial eta squared of .019, which is a small effect. These results indicate that there was a statistically significant difference among scores on the ATR for all 16 conditions (see Table 7).

Table 7.

Tests of Between-Subjects Effects for the Attitudes toward Research scale

Source	Type III Sum of Squares	*df*	Mean Square	F	Sig.	Partial Eta Sqaured
Back_Color	2520.732	1	2520.732	3.948	.048*	.019
Error	127701.831	200	638.509			
Total	5599706.00	216				

*Note: Significant at the $p<.05$ level.

Computer User Self-Efficacy

A 2x2x2x2 CRF-ANOVA was performed in order to determine whether there was a statistically significant interaction among font color, background color, highlight bar, and

progress bar indicator on responses to Part 2 of the CUSE scale. Analyses of the data for Part 2 of the CUSE identified a significant two-way interaction [F(1,195)=6.844,; *p*=.017] between font color and background color (see Table 8). The significant effect was associated with a partial eta squared of .034, which is a small effect.

Table 8.

Significant interaction of Font Color and Background Color for Part 2 of the Computer User Self-Efficacy scale

Source	Type III Sums of Squares	*df*	Mean Square	F	Sig.	Partial Eta Squared
Font x Back_Color	2820.503	1	2850.503	6.844	.017*	.034
Error	80361.763	195	412.112			
Total	89426.664	211				

*Note: Significant at the *p*<.05 level.

Because the interaction of font color and background color was significant, a test of the simple main effects was performed to reveal the nature of the interaction. Table 9 presents the test of the simple main effects for the interaction of font color and background color.

Table 9.

Tests of Between-Subjects Effects for Part 2 of the Computer User Self-Efficacy scale

Source	Type III Sums of Squares	*df*	Mean Square	F	Sig.	Partial Eta Squared
Font x Back_Color	2551.345	1	2551.345	6.129	.014*	.029
Error	86164.279	207	416.253			
Total	4092545.00	211				

*Note: Significant at the *p*<.05 level.

Table 10 shows the univariate F test of the simple main effects of 'Font' within each level combination of 'Background color.' For a white background, there were statistically significant differences among responses to Part 2 of the CUSE.

Table 10.

Univariate tests of the simple main effects for Part 2 of the Computer User Self-Efficacy scale

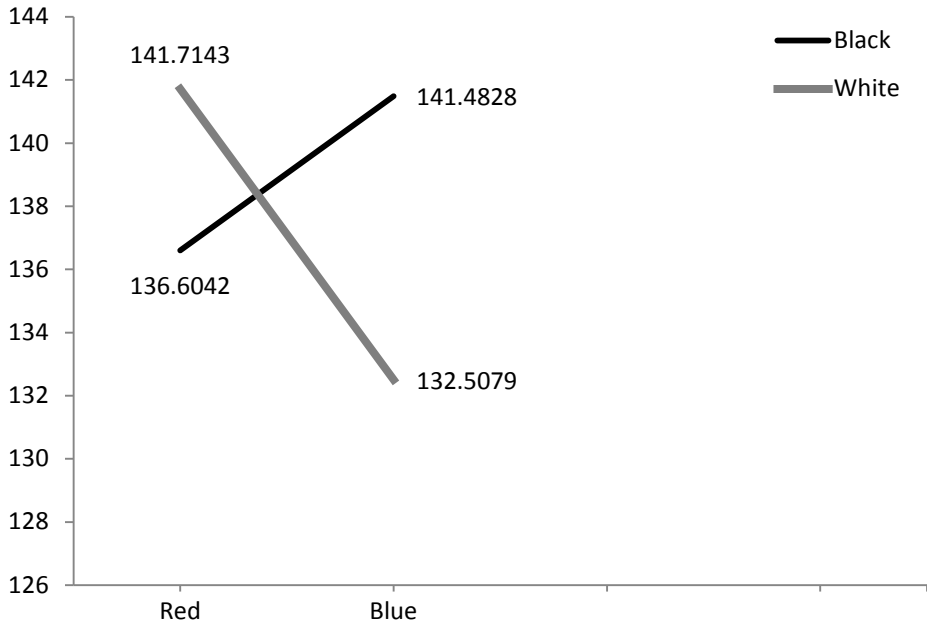
Back_color	Sums of Squares	*df*	Mean Square	F	Sig.	Partial Eta Squared
Black	625.104	1	625.104	1.502	.222	.007
White	2135.873	1	2135.873	5.131	.025*	.024

*Note: Significant at the $p<.05$ level

Each F tests the simple main effects of Font within each level of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

Tests of the simple main effects indicated that a white background produced a statistically significant effect across the two font colors—red and blue—whereas a black background was not significant across font colors. Figure 1 shows that for scores on Part 2 of the CUSE scale, significant differences were found for participants who responded to surveys with a white background versus a black background, regardless of font color. That is, participants who responded to a survey with a white background scored significantly higher on Part 2 of the CUSE scale than participants who responded to a survey condition in which a black background was present.

Figure 1. Estimated marginal means of CUSE Part 2



Qualitative analyses

Qualitative analysis of the data involved identifying and analyzing participants' text responses to an invitation to provide feedback regarding their experience during the survey. The text provided by participants was analyzed in order to get a better understanding of their thoughts about the entire survey process. While no formula exists to analyze textual data, Patton suggests the analysis is reflected in the method in which the data were collected. That is, the text provided by participants in this study served as a "reality" of the situation in which the participants were responding to the survey. In other words, the participants responded to the stimuli (i.e., interaction of font color, background color, highlight bar, and progress bar) presented to them during the survey process. Text responses are the unit of analyses because they are provided by the participants themselves and do not function as a result of predetermined hypotheses (see Schutt, 2012).

The analysis was guided by the question of whether or not a) font color, b) background color, c) use of a highlight bar, and d) use of a progress bar indicator were well- or poorly-received. Participants were allowed to provide text responses to the following question: "*Please*

provide any comments you may have regarding the survey process, including your thoughts on the visual design elements (e.g., color of font, color of background, use of the progress bar indicator, and use of the highlight bar) of the survey." At the end of the survey, participants were asked to provide their thoughts about the overall visual design characteristics of the survey in which they viewed. Response to this invitation was voluntary; participants did not have to provide feedback if they did not wish to do so. One-hundred-and-two participants provided feedback about their experiences.

Qualitative analysis of the textual data was inductive. It involved the identification of patterns in the "textual data" obtained from the participants' viewpoints and opinions of the design, or visual, characteristics in the survey condition in which they viewed. All of the statements, or text responses, provided by participants were combined in a single Excel spreadsheet. Although all of the statements appeared in a single location, they were separated, or blocked together, within the Excel spreadsheet according to the survey (testing) condition from which they were obtained. That is, text responses from testing condition one (X1) were grouped together separate from the text responses of the other testing conditions. Likewise, text responses from testing condition two (X2) were grouped together separate from the text responses of the other testing conditions, and so on and so forth.

Initial analyses explored the similarities and differences in the statements provided by different participants in the same testing condition. Text responses were examined for general themes or patterns within that testing condition. Using this data reduction strategy, the goal of the preliminary analyses was to determine which responses constituted "meaningful" or "relevant" data. That is, which responses best contributed to answering the research question(s). Analyses of the textual data also included looking for "atypical" responses or "deviation from" patterns (National Science Foundation, 1997). Statements were then organized and categorized into concepts resulting from specific expressions and phrases in the responses. General themes emerged from the textual data in accordance with the underlying research question. Such themes

included a request for more “standard” [font] color combinations” and/or a less “harsh” color for the highlight bar. Many participants stated they did not notice a progress bar or wished the progress bar “indicated a numerical percentage.” Furthermore, participants expressed the overall design scheme was “horrible,” “unattractive,” and “difficult to read at times.”

From there, analyses of the text responses were situated in how the interaction of these design characteristics was perceived. That is, statements were analyzed as a whole, instead of individual, or belonging to a certain group (i.e., testing condition). However, due to the nature of the interaction of font color, background color, highlight bar, and progress bar, the statements were interpreted in the situation, or context, in which they were obtained. Thus, analyses of the data using this method were more extensive than the initial analyses because it divided the data into several groups instead of a single general view of the elements of survey design.

Font color

Red. The red font was present in eight of the 16 conditions. Of the eight conditions in which the font color was red, four conditions presented red font against a black background while four conditions presented red font against a white background. Of the four conditions in which the red font was present, two conditions included a yellow highlight bar while two conditions did not (see below for a detailed outline of how the highlight bar functioned within the context of each of the testing conditions). Of the 102 total text responses, 44 were collected from conditions in which red font was present.

Analyses into the qualitative data obtained from participants showed a strong dislike for red font against both a black and white background. Red font, when presented against a black background with no highlight bar present was viewed as “very distracting” and “difficult to read at times.” Red font presented against a white background, with no highlight bar present, did not “provide enough contrast.” Other feedback indicated a dislike for “red font” and that “red font is somewhat harder to read.”

When presented against a yellow highlight bar, regardless of background color, red font was also viewed negatively. (In Qualtrics, the questions are initially presented against either a black or white background. When the participants click on a question, the text is 'highlighted' by the yellow highlight bar, as to focus their attention on the current question being answered. The text of the present question, then, is presented against a yellow background while all other questions continue to be seen against either a black or white background, depending on the participants' testing condition). For red text viewed against a yellow highlight bar, one participant stated that, "[while] the red font is appealing, the yellow is a little too bright for me. Another participant stated that the "survey was very hard to read." Other statements suggested the use of a "bright red with a bright yellow highlight" made the text significantly less visible.

Blue. The blue font was present in eight of the 16 conditions. Of the eight conditions in which the font color was blue, four conditions presented blue font against a black background while four conditions presented font against a white background. Of the four conditions in which the blue font was present, two conditions included a yellow highlight bar while two conditions did not. Of the 102 total text responses, 58 were collected from conditions in which blue font was present.

Analyses into the textual data concerning the blue font were considered difficult to read against a black background but received mostly positive feedback when presented against a white background. For testing conditions in which blue font was presented against a black background and no highlight bar was present, participants viewed the color scheme as "difficult to read" and "unprofessional." Blue font was perceived as "too dark" and "clashes with" the black background.

The blue font conditions were also viewed more positively when the highlight bar was present although the yellow highlight bar, in general, was not well received (see below for a general discussion and feedback concerning the presence or absence of the highlight bar). When blue font was presented against a yellow highlight bar, responses revealed that while "it was hard

to see with the black background and blue font, the highlighting helped" and that "without the highlight bar, this survey would have been nigh impossible to complete." However, one participant stated there was "way too much contrast and brightness with the highlight bar." Several other responses described the survey as "difficult" or "hard" to read.

Background color

Black. The black background was present in eight of the 16 conditions. Of the eight conditions in which the background was black, four conditions contained red font and four conditions contained blue font. Thus, participants either viewed red or blue font against a black background. Of the four conditions in which red font was present, two conditions included a yellow highlight bar while two conditions did not. This format is the same for conditions in which blue font was present. Of the 102 total text responses, 59 were collected from conditions in which a black background was present. Twenty-six responses were collected from conditions in which red font was presented while 33 responses were collected from conditions in which the font color was blue.

Feedback collected from participants regarding the black background revealed only two positive statements. For example, one participant stated that the dark background "[kept] my eyes from straining as much" while another participant stated that the "black makes the blue pop." Both of these statements were obtained from participants who viewed survey designs in which the highlight bar was present. In contrast, many participants described the blue and black color combination as "horrible" and "way too hard to read." For the red font with black background conditions, there was only one positive response.

White. The white background was present in eight of the 16 conditions. Of the eight conditions in which the background was white, four conditions contained red font and four conditions contained blue font. Thus, participants either viewed red or blue font against a white background. Of the four conditions in which red font was present, two conditions included a yellow highlight bar while two conditions did not. This format is the same for conditions in which

blue font was present. Of the 102 total text responses, 43 were collected from conditions in which a white background was present. Thirteen responses were collected from conditions in which red font was presented while 25 responses were collected from conditions in which the font color was blue.

Analysis of the textual data regarding the use of a white background yielded negative results. Red text on a white background, regardless of the presence or absence of the yellow highlight bar, appeared to most participants as "garish," "jarring," and "not the easiest to read." For conditions in which blue font and the yellow highlight bar was present, participants viewed the color scheme as "obnoxious" and "distracting." Only one (n=1) participant considered the red font-white background color scheme favorably, stating that it was "better than the normal black." For the blue font-white background color combination, two (n=2) participants provided positive feedback: "the blue font is easier to read than black font and is visually more appealing than black font" and "[the] word color with [the] background color made the words easy to read."

Highlight bar

The highlight bar was present in eight of the 16 testing conditions. As stated before, the purpose of the highlight bar was to help focus the participants' attention to one question at a time. When participants clicked on a question, the highlight bar "highlighted" the question and participants viewed either red or blue text against a yellow "background." Therefore, feedback concerning how participants viewed either red or blue text is perceived in this context.

Of the eight conditions in which the highlight bar was present, 53 responses were recorded. Out of the 53 responses, 19 included feedback regarding the use of the yellow highlight bar. For all conditions in which the highlight bar was present, the majority of participants stated the color of the highlight bar was "too bright" or "painful to the eyes." However, in some conditions, participants indicated the highlight bar "helped improve focus on one question at a time" and "not get lost in all of the questions."

Progress bar

The progress bar was used as an indicator for how much of the survey participants had completed and in turn, how much they had not completed. In Qualtrics, the progress bar is a small rectangular-shaped icon at the bottom of each page of the survey. The image displays 0% and 100% on either side of the icon and "Survey Completion" above the progress bar to give further details to the respondent about their completion status (Qualtrics, 2014). The image does not, however, show the exact completion percentage during the process of completing the survey.

The progress bar was present in eight of the 16 conditions. Analyses into the textual data (n=11) provided by participants revealed the progress bar indicator was shown to be only "mildly informative," with most participants stating they wished it provided more information such as a "numerical percentage." Many participants indicated that they did not even notice the progress bar, with some participants stating that it was "rather small" and [seemingly] not very important." One participant stated that the progress bar is a "good feedback device" while another revealed that it helped to give "a feeling of accomplishment."

Summary of responses

The table below presents a summary of the feedback obtained from participants about their experiences responding to their assigned survey. Not all textual data are provided here. (For a complete list of feedback, see Appendix E). Statements included in Table 9 represent the general consensus of participants who responded to the request for feedback regarding the visual (design) characteristics of the survey in which they viewed.

Table 11.

Summary of responses from participants

Group number	Testing condition	Statement/feedback
1	Red font; black background; highlight bar present; progress bar present	"…the colors need to be [simpler]. Just black and white would be better than red and yellow."

		"The dark background keeps my eyes from straining as much and the red font is appealing."
2	Red font; black background; highlight bar absent; progress bar present	"The red font is distracting." "The red and black color scheme made the overall presentation very serious…"
3	Blue font; black background; highlight bar present; progress bar present	"Do not like the highlight bar." "The blue on black text is difficult to read."
4	Blue font; black background; highlight bar absent; progress bar present	"I haven't seen blue font used on surveys much." "…it makes it difficult to read in comparison to the usual white/black text."
5	Red font; black background; highlight bar present; progress bar absent	"There wasn't much difficulty reading the red text over [the yellow]." "The colors are interesting."
6	Red font; black background; highlight bar absent; progress bar absent	"[The] red font was terrible." "The black and red are awful color choices for a survey."
7	Blue font; black background; highlight bar present; progress bar absent	"Difficult to read." "With [the highlight bar], a combination of blue text on a yellow background made it somewhat uncomfortable to view the questions and answers."
8	Blue font; black background; highlight bar absent; progress bar absent	"I would never use this scheme." "Visual elements definitely need to be changed."
9	Red font;	"The highlight bar is ok[ay],

	white background; highlight bar present; progress bar present	as well as the status bar." "Color of font and highlight annoyed me greatly."
10	Red font; white background; highlight bar absent; progress bar present	"The red font is difficult to read." "It would be nice if the progress bar indicated what numerical percentage had been completed."
11	Blue font; white background; highlight bar present; progress bar present	"The font color didn't seem to have any effect on me." "The blue of the font bothers my eyes."
12	Blue font; white background; highlight bar absent; progress bar present	"Color of font is difficult to read." "Why not make [the font] black?"
13	Red font; white background; highlight bar present; progress bar absent	"I'm not a fan of the red font…" "…it was better than the normal dull black."
14	Red font; white background; highlight bar absent; progress bar absent	"I don't like [the red]." "Not a fan of [the] red font."
15	Blue font; white background; highlight bar present; progress bar absent	"The blue font is… visually more appealing." "The blue font was easy to read."
16	Blue font; white background; highlight bar absent; progress bar absent	"[I] liked the blue; it helps my eyes." "The font color was tiring for my eyes."

Response rate

The analysis of data under the response rate condition was guided by the question of whether or not the visual characteristics of the survey influenced the rate at which participants completed the survey. That is, whether or not the interaction of a) font color, b) background color, c) use of a highlight bar, and d) use of a progress bar indicator affected the response rate for each of the 16 testing conditions. Initial examinations of the response rate—the rate of completion of the survey questionnaire—for each of the 16 conditions showed a range of 53.8 percent to 86.4 percent, with a mean of 74.12 percent. Descriptive statistics for response rate can be seen in Table 12.

Table 12

Descriptive statistics for response rate

Mean	74.118
Standard deviation	10.636
Minimum	53.80
Maximum	86.40
Range	32.60
Count (N)	16

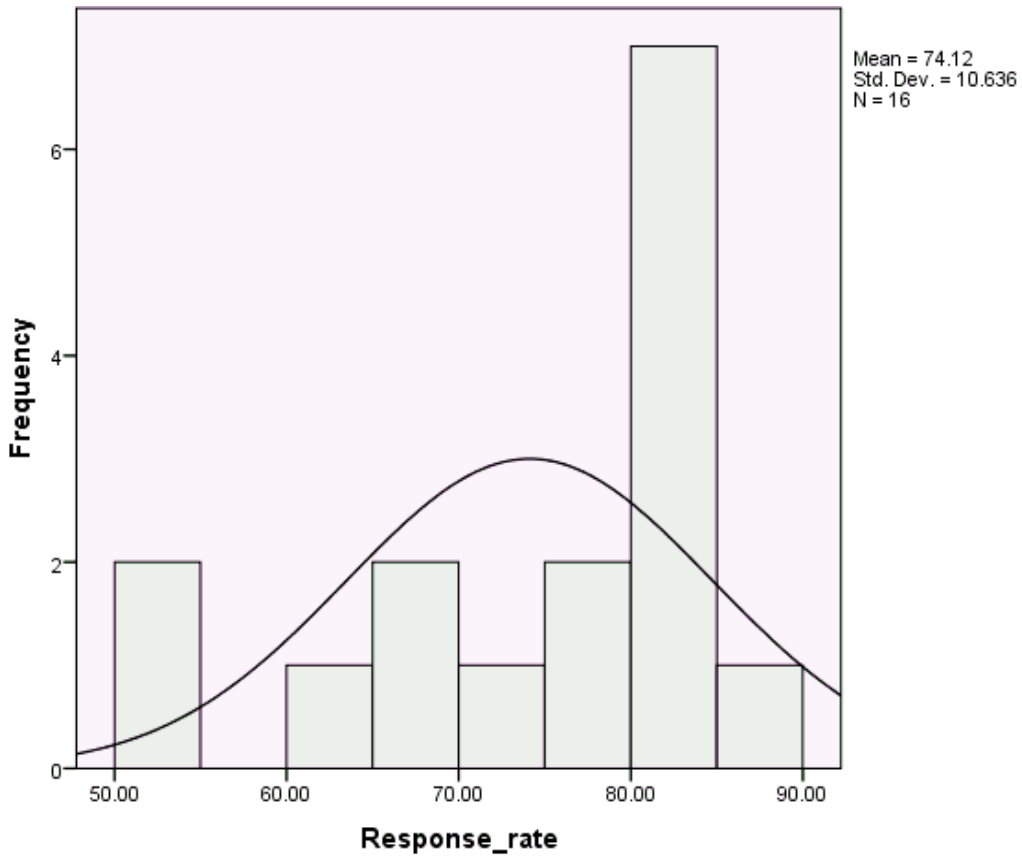
Statistical analyses of the response rate showed that for conditions in which extremely negative feedback was obtained, the response rate was 53.8 percent. As Table 13 shows, two conditions had a response rate of 53.8 percent. These conditions—groups 1 and 13—were characterized by red font with the presence of a yellow highlight bar. Two conditions—groups 2 and 8—had a response rate of 81.8 percent. For each of these conditions, the yellow highlight bar was not present. The highest response rate was found for the condition—group 15—in which feedback obtained from participants was more favorable (see Table 10). When blue font was

presented against a white background with the yellow highlight bar present, the response rate was 86.4 percent.

In order to compare the mean response rates from different groups of data, z-scores were used. (A *z*-score is a standard score that allows data obtained from different groups to be compared. Data are standardized so that they can be reflected on the same scale, i.e., the normal distribution. A *z*-score reflects the number of standard deviations an observation or datum obtained from each group is above or below the sample mean). Figure 2 shows the distribution of the group mean response rates. The spread of scores is reflected by the size of the standard deviation (*SD*= 10.636) and produces a slightly negative skew. This indicates that most of the average response rates were above the mean of 74.12. Specifically, eight groups (i.e., conditions) of the 16 total conditions had a response rate of 80.0 percent or more.

As Figure 2 shows, two conditions have response rates nearly two standard deviations from the mean (*M* = 74.12).

Figure 2. Distribution of group mean response rates



In order to identify outliers—observations that appear to deviate from other observations in the sample—*z* scores were used. Table 13 shows the response rate for each condition and its corresponding *z* score.

Table 13.

Z scores for the response rate variable

Condition (group)	Response rate (percent)	Z score
1	53.8	-1.910*
2	81.8	0.722
3	61.1	-1.224
4	70.0	-0.387
5	78.6	0.421

6	77.8	0.346
7	80.0	0.553
8	81.8	0.722
9	82.4	0.779
10	83.3	0.863
11	68.4	-0.538
12	80.0	0.553
13	53.8	-1.910*
14	81.3	0.675
15	86.4	1.155
16	65.4	-0.820

Summary of results

The pages presented in this chapter discussed the results of the quantitative and qualitative analyses guided by the research questions. Analysis of the data collected from 216 participants revealed a significant two-way interaction on the Part Two of the Computer User Self-Efficacy Scale between font (red or blue) and background color (black or white; F[.05,1] = 6.844, p=.01). Tests of the simple main effects showed a white background resulted in a statistically significant effect across the two font colors, red and blue. Analysis of the data also revealed a main effect for background color—black or white—on the Attitudes toward Research measure. In contrast, no significant interaction among the variables was obtained for the Knowledge instrument or Completion Time. (Means and standard deviations for the Government Knowledge test, 'Attitudes toward Research' scale, and 'Computer User Self Efficacy- Part 2' scale are presented in Appendices F-I, respectively).

Qualitative analyses showed a strong negative response to red font, regardless of the background color. When presented against a yellow background, participants seemed to reveal an

even stronger aversion to the color combination. This aversion was also apparent in conditions in which blue font was presented against a black background with no highlight bar present. Blue font was most positively regarded when presented against a white background. The use of a highlight bar proved to be only somewhat helpful in helping participants focus on one question at a time. The general consensus of participants who responded to the request for feedback concerning the design elements of their particular survey revealed a preference for standard color combinations, such as darker font on a lighter background (e.g., black text on a white background).

CHAPTER V

DISCUSSION

The purpose of this chapter is to present a discussion of the results obtained through analysis of the data. The chapter is divided into six sections. The first section presents a review of the study. The second section summarizes the major findings of the study and is followed by a section discussing the relevance of these findings. The fourth section suggests implications for future research based on the findings presented in Chapter IV. The fifth section discusses the limitations of the study. The sixth and final section presents conclusions drawn from the study.

Review of the study

The purpose of this study was to explore the effects of variations in the combinations of the independent variables—font color, background color, highlight bar, progress bar indicator—on participants' responses to questions presented in an online format. These combinations were presented in 16 different online survey conditions (these combinations are more fully described in Chapter III, Table 1). The study was guided by the following research questions:

1. Does text color, background color, appearance of progress bar, or appearance of highlight bar affect the a) response rate, b) completion time, and group scores on c) a knowledge test, d) attitudinal measure, and e) self-efficacy scale when administered in a web-based mode?
2. Are there significant interactions of text color, background color, appearance of progress bar, or appearance of highlight bar on a) response rate, b) completion time, and group scores on c) a knowledge test, d) attitudinal measure, and e) self-efficacy scale when administered in a web-based mode?

Summary of the major findings

As stated in Chapter IV, four separate 2x2x2x2 completely randomized factorial analysis of variances—one for each of the three instruments and one for completion time variable—were conducted in order to determine the interactive effects of font color, background color, highlight bar, and progress bar indicator on participants responses to survey questions administered via the Internet. Differences among groups for the response rate variable were determined using z-scores. A summary of the findings are as follows:

Analysis of the data collected from 216 participants showed no significant main effects or interactions among text color, background color, appearance of progress bar, or appearance of highlight bar on answers to Government knowledge questions.

Analysis of the data revealed a statistically significant main effect for background color on responses to the Attitudes toward Research scale [$F(1, 200)= 3.948$; $p=.048$] for all 16 groups.

Analysis of the data also showed a statistically significant interaction [$F(1,195)=6.844$; $p=.017$] between font color and background color on responses to Part 2 of the Computer User Self-Efficacy scale. Tests of the simple main effects revealed that a white background resulted in a statistically significant effect [$F(1,207)=6.129$; $p=.014$] across the two font colors, red and blue, whereas a black background resulted in a non-significant effect. Significant differences ($p=.025$) among groups were found for conditions in which a white background was present.

No significant main effects or interactions were found for completion time.

For the response rate variable, two groups had response rates approximately two standard deviations from the mean (74.12).

Discussion of the findings

The use of online technologies to administer surveys has increasingly pervaded the field of survey research. Specifically, the expanding availability of web-based survey software programs has allowed researchers to easily and rapidly create online questionnaires to gather data from large numbers of respondents. Compared to paper based surveys, online administration allows researchers the freedom to make the survey "more attractive" through the use of visual design aspects not readily available through traditional paper-and-pencil administration modes. This study implored Qualtrics, an online survey tool available for free to OSU researchers and students. Qualtrics offers numerous preset "themes" that contain default colors, fonts, and other stylistic elements. Survey researchers also have the option to "customize" their survey using the program's countless design element options (i.e., font type, color, and other advanced style choices). The surveys included in this study were constructed using custom colors (for both font and background) as well as other visual effects offered in Qualtrics' advanced options (i.e., highlight bar and progress bar indicator).

This study implored a mixed mode strategy in which different forms (N=16) of the same online survey were administered to groups. As detailed in Chapter I, a mixed mode methodology allows researchers to determine the effects of mode on participants' responses. Previous studies have often implored mixed mode methodologies in which different strategies, including sampling procedures, were used to compare responses between modes. These strategies often lead to bias (i.e., mode effects; Vannieuwenhuyze & Loosveldt, 2012) which reduces the comparability of survey modes. The study presented here sought to diminish those biases by administering different forms of the same survey in the same mode (i.e., online). More precisely, the study was designed to identify the effects of certain visual design elements by administering different forms of the same survey.

While the surveys contained the same verbal information (i.e., survey content was identical), they differed in terms of font color, background color, use of a highlight bar, and use of a progress bar indicator. In order to ensure data quality, a single sample of participants was selected to respond to the surveys. Unlike mixed mode methodologies that use different samples of respondents across modes, a single sample of respondents was recruited to participate in the study. This sampling method allowed for more strict comparisons among the survey modes.

The significant findings revealed in Chapter IV provided support for previous studies (Best & Krueger, 2004; Brock, Barry, Lawrence, Dey, & Rolffs, 2010; Stern, Smyth, & Mendez, 2012) that found the use of color impacted participants responses to questions presented in an online format. In the current study assessing the effects of color (i.e., background color, *black versus white*) in online surveys, responses to the Attitudes toward Research scale were statistically significantly different among the 16 groups. This finding suggests that background color plays an important role in the response process. However, the analysis did not reveal which background color produced the significant effect.

It is important to note that background color is inevitably linked to and perceived only in comparison to font color. The surveys in this study presented either a black or white background in combination with either red or black text. Based on participant responses collected in feedback portion of the surveys it can be assumed participants regarded a white background more favorably, as there were more positive responses for conditions in which a white background was present (*see Table 12*). Basic color theory as well as basic principles for survey design recommends white (or light) backgrounds is best for web-based pages. Conceivably, this is because the color white offers a high contrast to a wide range of colors including, but not limited to, red and blue.

Previous research has shown that participants reacted negatively toward black and white surveys and advocated for use of “non-traditional” color combinations. While Hill and Scharff (1999) proposed the use of contrasting colors and asserted this color combination approach assists

readability, Ling and van Schaik (2002) promoted the use of contrasting colors (e.g., white/black, blue/yellow, red/green) in order to enhance performance and increase legibility. Thus, the visual appearance of a survey is best when dark text presented on light background or when "visually appealing" color combinations are used. According to color theory for web design, the use of contrasting colors helps prevent eyestrain and focuses readers' attention on specific page elements. The most apparent use of contrasting colors is found through the link between font color and background color. Specifically, when dark text is presented against a light background or when light text is presented against a dark background, the text is easier to read and draws the eyes to the main content of the page.

Regarding the link between font color and background color, the significant findings on Part 2 of the CUSE revealed the importance of combining colors on web-based surveys. A significant interaction was revealed between font color and background color. This finding suggests that responses to Part 2 of the CUSE were significantly different among the 16 groups due to the influence of font color and background color. That is, the color combination of red text on either a black or white background and blue text on either a black or white background affected participants' responses. The colors chosen for this study were selected based on their high level of contrast. A test of the simple main effects for the interaction of font color and background color revealed that for conditions in which the background was white, responses to Part 2 of the CUSE were statistically significantly different than conditions in which the background was black. Participant feedback seems to support this finding. More positive reactions to the color combinations of red/white and blue/white were found, especially when the yellow highlight bar was absent.

Color theory supports the use of the red/black and red/white color combination as well as the blue/white and blue/black color combination. It is important to note that the yellow highlight bar is embroiled within the font/background color combinations for conditions in which the highlight bar was present. The color yellow, according to color theory, is opposite to blue on the

color wheel and therefore the two colors are considered complementary to the each other. When used together, these colors should enhance one another. However, no significant interaction was found for font color (red or blue), background color (black or white), and use of the yellow highlight bar. Therefore, it is uncertain how much of a role the highlight bar influenced responses.

The purpose of the highlight bar was to focus the participants' attention to one question at a time. Yellow as chosen as the highlight bar color based on its high contrast to both red and blue. Effective text legibility occurs when the font color is highly contrasted to the background color. Additionally, many proponents for the use of color on web page design maintain the use of contrasting colors to improve text legibility. However, narrative feedback from participants about the color and design scheme of the surveys in this study suggested the color combinations were too highly contrasted and did not, in fact, make the survey easier to read. Specifically, many participants expressed a strong disdain for the red/black color combination especially when the yellow highlight bar was present. For conditions in which blue font was presented against a black background and the highlight bar was present, participants stated the visual appearance of the survey was "horrible" and "unattractive." Participants' feedback for conditions in which a white background was present and included the highlight bar also seemed to suggest an aversion to the yellow highlight bar.

Lighthouse International (2014), a pioneer in the field of vision rehabilitation, provides specific guidelines for making effective color choices that work for people with and without vision deficiencies. Additionally, WebAIM (2014a) offers a 'Color Contrast Checker' that allows web page designers to test the "contrast ratio" of foreground color (e.g., font) to background color. This tool serves as a way to verify the "accessibility" of the web page and ensure the colors are sufficiently contrasted to one another. Like Lighthouse International, WebAIM (2014b) seeks to ensure that web pages are usable to persons with disabilities. This is especially important given that a few of the participants in this study expressed they had visual impairments that made completing the survey extremely difficult, if not impossible.

Nearly all of the participants in the current study expressed a strong aversion to the colors. This is somewhat surprising in that each of the colors in the study were of high contrast to all the other colors in the survey, as previous literature has shown this strategy to be highly effective. However, when the WebAIM Color Contrast tool was used to verify the level, or effectiveness, of the contrasts, several of the color combinations failed to meet web accessibility standards. (For the purposes of clarifying the effect of the colors used in this study, the following information is pertinent to the understanding of the importance of effective color contrasts):

> "WCAG 2.0 consists of three priority levels that act as an industry standard. The first level, Level A, covers items on web pages that must be made accessible in order for individuals with disabilities to access the content at all. The second level, Level AA, includes items on web pages that should be made accessible to allow a wider group of users to access the content. Level AAA describes items on web pages that can be made accessible to allow the widest amount of individuals with disabilities to use the site."

In particular, the red font/black background combination when coupled with the yellow highlight bar did not pass the guidelines for effective color contrast when using 12-point font. Similarly, 12-point red font on a white background combined with the yellow highlight did also not meet web content for accessibility standards. This suggests that red and yellow are not complementary and should not be used in combination with one another. Narrative data seems to support this assertion as the majority of respondents found this color scheme made the survey difficult to read and displeasing to the eye.

Participant feedback from groups where blue font was presented against a black background revealed great difficulty in the ability to read the survey questions. This is not surprising in that the combination of blue and black failed all levels of the web content accessibility guidelines (WCAG). Likewise, normal (i.e., 12-point) red font on a white background failed both Level AA and AAA guidelines. When 12-point red font is observed against a black background, the color contrast ratio also failed to meet Level AAA standards (though it did meet WCAG Level AA). Given this, it can be suggested that this color scheme

should not be used in survey design. The fact that no significant effect was found for either font color on any of the three instruments is somewhat surprising. Red font, in general, produced more negative feedback while blue font was generally well-received.

The blue font/white background color combination is generally considered to be a highly effective color contrast (blue/white passed all levels of the web content accessibility guidelines). However, the nine responses collected from participants who experienced these conditions revealed very little acceptance of this design scheme. Similarly, blue font on a white background with the yellow highlight bar and blue font on a black background with the yellow highlight bar passed all levels of the accessibility guidelines. Again, feedback was not positive. Thus, taken together, the narrative data (i.e., feedback) for these conditions is surprising. According to WebAIM's WCAG criteria, the blue and white and blue and yellow are effective color contrasts and should be well-received. That is, text should be highly legible. The fact that feedback revealed that less contrast would be preferable warrants more investigation. One possible explanation is that participants are not accustomed to seeing "uncommon" color combinations, despite their self-reported familiarity with the computers and the Internet, and anything other than standard black and white requires more visual processing.

The absence of any significant effect for text color, background color, appearance of progress bar, or appearance of highlight bar on responses to the Government knowledge test is noteworthy. Specifically, the Attitudes toward Research scale and the Computer User Self-Efficacy scale are self-report measures that gather data using a Likert scale while the Knowledge test contained items with one correct answer. It is plausible that the nature, or content, of the test was different in such way that responses could not be influenced regardless of the design elements. In particular, Part 2 of the Computer User Self-Efficacy scale measures participants' self-reported ability to effectively use computers. While this scale does not specifically address proficiency with the web-based programs, it is plausible that participants' familiarity with the

Internet has contributed to their exposure to numerous design schemes including various color combinations.

Survey completion time was also not significantly influenced by the surveys' visual characteristics. Given participants responses revealed that the design schemes of the surveys made the text difficult to read, this finding is particularly interesting. More specifically, feedback from participants in each of the 16 groups was nearly identical with participants expressing a similar strong aversion to each of the design schemes. Taken together, the narrative data collected from participants suggests that the designs elements used in this study were not well-received. The fact that completion times did not statistically significantly differ among groups suggests there may be an inherent similarity among unconventional design schemes. That is, it is conceivable that the design schemes were, in and of themselves, equivalent.

Analyses of the response rates among the 16 groups showed that for two conditions in which red font was present and coupled with a yellow highlight bar, the response rate was 53.8 percent. After converting the response rate percentages to z-scores, which allowed for a more strict comparison of group response rates, it was apparent a response rate of 53.8 percent was located two standard deviations from the study's overall mean response rate of 74.12 percent. This suggests that certain elements of these two survey conditions were in some way different than the other conditions. Based on participant feedback, it can be assumed the strong aversion of the red font and yellow highlight bar contributed to this low percentage. Further support for this claim can be found in the guidelines for web content accessibility (i.e., the combination of red and yellow failed to meet any levels of the WCAG criteria for effective color contrast).

This study sought to consider the effects of color on participants responses to items presented in an online format. Schaeffer and Dykema (2011) contend stylistic elements such as color as "nonverbal information" inevitably contributes to how participants respond to web content. Careful consideration was taken when selecting the colors used in this study. Specifically, previous literature informed the decision to use high contrasting color combinations.

However, textual data obtained from participants during the feedback portion of the surveys revealed extremely negative assessments of the color choices. Given this, it is highly plausible that the color combinations were too extreme. According to color theorists, the more vibrant the colors the more mental energy that is expended.

While the use of color—font and background—in surveys is often based on researchers personal preferences, the aesthetic quality of particular colors and/or color combinations must be cautiously executed. Some color combinations are more favorable than others, such as blue and yellow or blue and white, while others should be avoided (e.g., red and white, red and yellow). Additionally, some design aspects offered in online survey programs, such as a progress bar indicator, do not seem of great significance. The results of this study indicated progress bar indicators are only mildly informative and do not seem to influence responses. Additionally, while the highlight bar did not produce a significant effect for any of the 16 groups, it appeared to achieve a purpose. For conditions in which it was present, participants stated that it helped them to focus on one question at a time. However when coupled with red font, the highlight bar was considered “too bright.”

Collectively, the findings of this study provided valuable information concerning the use of color in online surveys. Discernibly, effective color contrasts must be achieved in order to meet certain guidelines that allow web-based pages to be legible for nearly everyone. The unexpected issue of participants who expressed concern over their ability to participate in the study due to visual deficiencies served as the basis for the discussion presented in this section. Additionally, participant feedback played a pivotal role in the interpretation of how the study’s design schemes were received by the participants. While the majority of feedback was negative and indicated a preference for more traditional (i.e., black and white) color combinations, it is crucial researchers continue to consider the use of effective color contrasts when designing surveys.

Limitations of the study

This study was limited by several key factors. First, the restrictions placed on OSU IRIM for recruiting students for participation in research limits the number of times students can be contacted. Students were sent an invitation to participate on two separate occasions, one three weeks after the start of the Fall 2013 semester and one approximately three weeks after the initial request. While this is not necessarily a limitation, it is possible more participants could have been recruited given that additional invitations were allowed. Also, it is important to note that the data collection period for this study occurred during the course of six weeks (i.e., September-October). Given a longer timeframe, it is highly plausible more data could have contributed greatly to the overall results of this study. Additionally, the invited sample of 4,800 students resulted in a study sample size of 216 participants, an overall response rate of 4.5 percent. Taken together, these limitations contributed to a small sample size.

Another limitation of this study is reflected in the characteristics of the sample. All participants were sampled from a single population (i.e., OSU) of university students and do not necessarily reflect the views of the general population. Likewise, although best efforts were made to obtain a random sample of OSU students, it is not guaranteed that the study sample adequately represents all OSU students. For instance, while specific ages of participants were not obtained, the majority of participants (74.5%) were 18 to 24 years of age. Thus, the study did not include a substantial number of older adults who may respond significantly different than their younger counterparts.

Furthermore, narrative data collected from participants revealed a number of unexpected issues arising from the use of the Qualtrics software. First, participants appeared to experience problems when scrolling down the survey page. Second, participant feedback indicated that for conditions in which the background was designed to be black, the background was viewed as graduated gray and black. Third, the physical environment (i.e., bad lighting, glare on screen, etc.) and visual settings (i.e., low quality monitors, computer display settings, etc.) in which

participants responded to the survey were varied. That is, different visual displays (i.e., computer settings such as screen size, resolution, and browser settings; Brock, Barry, Lawrence, Dey, & Rolffs, 2010) may require different strategies for interpreting items due to the various computer processors (e.g., Windows, Mac, mobile phones, etc.) used for web-based surveys. Additionally, some participants expressed vision deficiencies that made responding to the survey difficult. Taken together, these unanticipated issues reflect the shortcomings of the design capabilities of online survey software programs.

The instruments used in the study comprised a total of 93 questions. This is a relatively high number of items for a survey, especially surveys administered via the Internet. While participants were informed the study would take approximately 20 minutes to complete, the number and nature of the questions was not revealed. The content of the instruments, specifically the 'Attitudes toward Research' scale and the 'Computer User Self-Efficacy' scale, likely influenced how participants responded to items on these measures. That is, these instruments, although not correlated, were perceived by participants to possibly influence their responses. Furthermore, many participants indicated that they were uncertain about answers to many of the Government knowledge questions. It is possible that given different instruments, perhaps one with fewer questions, the results of this study could have been vastly different.

In regards to the use of color and other design elements, respondents were notified that they would be participating in a study that would extend the literature on the effects of variations concerning the visual presentation of online survey designs. Specifically, participants were informed that the research study would investigate how the interaction of text and background color affects how information appears visually to respondents, especially in an online format. Participants were not informed what colors or specific design elements would be used. It is possible that, upon viewing the color and design, participants reacted negatively toward the visual design and ceased participation in the study.

Qualtrics offers all colors offered by the RGB color model. Although best efforts were employed to make the colors used in this study as contrasted as possible, narrative data revealed the preference for simpler, more "traditional" color choices. The complex nature of the design elements likely contributed to a high non-response rate. That is, given different color contrasts, participants may have reacted very differently and completed the survey in its entirety.

Implications for future research

Section 508 of the Rehabilitation Act of 1973 (29 U.S.C. § 794d) requires the federal government to ensure that the electronic and information technology that it develops, procures, maintains, or uses is accessible to persons with disabilities. While Dillman, Tortora, Conradt, and Bowker (1998) found that non-completion is lower for surveys with elaborate features, Vincente and Reis (2002) maintain careful consideration must be taken when implementing visual enhancements of online surveys. This study provides support for the use of simpler design schemes, especially color choice. Participant feedback revealed the colors used in this study, especially the yellow highlight bar, were too vibrant and produced too much contrast with one another.

Taken together, the results of this study recommend future studies should bear in mind elaborate design schemes and explore more color combinations, such as light text on a dark background or dark text on a light background. Other design elements, such as larger text, may also be highly accommodating, especially given that physical environment (i.e., bad lighting, glare on screen, etc.) and visual settings (i.e., low quality monitors, computer display settings, etc.) often vary among participants. The availability of the WebAIM Color Contrast tool allows researchers to effectively design surveys for use on the web. Particularly, the tool ensures all individuals who may respond to requests to participate are able to view the survey with little or no difficulty. Future research should utilize this tool to ensure proper requirements are met to meet accessibility standards.

Additionally, the vast capabilities of web-based survey software package to create, design, and administer surveys have encouraged researchers to forego "traditional" paper-based modes. As stated in Chapter I, surveys administered via the Internet allow researchers to collect data more efficiently in terms of time, cost, and labor. However, while web-based surveys do provide considerable advantages over their paper-based counterparts, administering surveys via the Internet produces less control over the physical environments and visual displays of the participants in the study. That is, administering surveys via the Internet yields less control over how participants view the survey as participants often employ different devices (e.g., laptop, desktop, mobile phone, etc.), operating systems, (e.g., Android, iOS, etc.), Internet browsers (Firefox, Google Chrome, etc.,), and computer display settings (e.g., monitor size, resolution, Internet browser, etc.) to respond to surveys.

Similar research has studied the effects of physical environment and visual displays on participants responses to surveys administered via the Internet. While Barenboym, Wurm, and Cano (2010) found that responses to an online survey could be significantly affected by the mode in which participants completed the survey questionnaire (i.e., at a time and location that was convenient to the participant and using a computer of the participants' choice as opposed to completing the survey at a university computer laboratory setting wherein the time and location were predetermined and set up by the researchers and using a university-owned computer located in the university's psychology computer lab), it could not be determined whether the significant findings were due to mode effects (characteristics of the survey mode or testing conditions) or person effects (using different participants for the two survey modes). The present study explored the consequences of administering multiple surveys in the same mode—the Internet—and found mode effects occurred.

Like the Barenboym, Wurm, and Cano (2010) study, the present study employed different participants for the each of the survey "modes." (For the purposes of this study, 'modes' refers to the 16 different testing conditions in which the survey questionnaire content was the

same but the design elements were different; likewise, 'mode effects' refers to the influence of the surveys' design characteristics on participants' responses to the survey questionnaire items.). In the current study, the significance of color on participants' responses suggests mode effects may be due to the variation of the design characteristics of the survey, the physical environments of the participants, the display settings of the participants' devices, or combinations thereof. While the present study explored the effects of color on participants responses to questionnaire items presented in an online format, the unanticipated effects of the participants' physical environments necessitates future investigation. More specifically, future research should investigate the influence of color and other visual design elements of online surveys wherein the physical environments and computer display settings are identical.

Conclusions

Whereas previous literature (Daley, McDermott, Brown, & Kittleson, 2003) has encouraged survey designers to use "non-traditional" color schemes instead of the "traditional" black-and-white color combination typically found in paper-and-pencil modes, no known study exists that compares multiple online modes of administration. This study sought to generate future research on this topic. Generally, participants revealed a strong dislike for the survey colors selected for this study. Although the color combinations used in the study are of high contrast to one another, participant feedback suggested the color combinations were "too bright" and made the text "harder to read." The yellow highlight bar was also not well received but seemed to serve its intended purpose of compelling participants to focus on one question at a time. Based on participants' statements regarding the highlight bar, it seems the brightness of the yellow provided too much contrast to both the red and blue font.

According to Lighthouse International, web page designers must bear in mind the audience to which their web page will be made available. Certainly, this includes surveys administered via the Internet. As online survey programs become increasing prevalent among researchers, special attention must be paid to the use of color and other visual aspects of online

survey design. Researchers must consider how their survey will be administered and be prepared for any possible issues (e.g., visual deficiencies of participants) that may arise. That is, the use of color and other visual design elements must be carefully selected when designing online pages and executed in accordance to Section 508 of the Rehabilitation Act of 1973 (WebAIM, 2014c). WebAIM provides thorough web accessibility training as well as technical assistance for those looking to design highly sophisticated yet usable websites and web pages. The 'Color Contrast Checker' offered by WebAIM allows designers to test web content for accessibility ensuring that sufficient contrast is provided for all persons (i.e., those with varying degrees of disabilities and those without disabilities).

Future research should pay careful attention to the stylistic elements of online surveys. It can be assumed that most surveys are designed using basic, or standard, color combinations (e.g., black, or dark, font on a white, or light, background). While this color scheme is not inferior to other design schemes, based on the results of the current study, it is conceivable that responses could be influenced by more "sophisticated" or stylistic design schemes. Nevertheless, most survey researchers can agree that surveys should be designed to be aesthetically pleasing and incorporate color in a way that makes the survey interesting and enjoyable while producing salient, quality data. Furthermore, researchers should also fully consider the environment in which participants may be responding to and completing the survey questionnaire given that unanticipated mode effects may occur and bias results. Taken together, the results of this study provide considerable evidence for the influence of color in online survey design.

REFERENCES

Alessi, E. J., & Martin, J. I. (2010). Conducting an Internet-based survey: Benefits, pitfalls, and lessons learned. *Social Work Research, 34*(2), 122-128.

Bandura, A. (1978). Self-efficacy: Toward a unifying theory of behavioral change. *Advanced Behavioral Research Therapy, 1*, 139-161.

Bandura, A. (1982). The assessment and predictive generality of self-percepts of efficacy. *Journal of Behavior Therapy and Experimental Psychiatry*, *13*(3), 195-199.

Bandura, A., & Wood, R. (1989). Effect of perceived controllability and performance standards on self-regulation of complex decision making. *Journal of Personality and Social Psychology, 56*(5), 805-814.

Barenboym, D. A., Wurm, L. H., & Cano, A. (2010). A comparison of stimulus ratings made online and in person: Gender and method effects. *Behavior Research Methods, 42*(1), 273-285.

Bertot, J. C., & McClure, C. R. (1996). Electronic surveys: Methodological implications for using the World Wide Web to collect survey data. *Proceedings of the ASIS annual meeting, 33*, 173-185.

Best, S. J., & Krueger, B. S. (2004). Internet data collection. Thousand Oaks, CA: Sage Publications.

Birnbaum, M. H. (2004). Human research and data collection via the Internet. *Annual Review of Psychology, 55,* 803-832.

Booth-Kewley, S., Larson, G. E., & Miyoshi, D. K. (2007). Social desirability effects on computerized and paper-and-pencil questionnaires. *Computers in Human Behavior, 23,* 463-477.

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health, 27*(3), 281-291.

Brannon, R. (1981). Current methodological issues in paper-and-pencil measuring instruments. *Psychology of Women Quarterly, 5*(4), 618-628.

Brock, R. L., Barry, R. A., Lawrence, E., Dey, J., & Rolffs, J. (2010). Internet administration of paper-and-pencil questionnaires used in couple research: Assessing psychometric equivalence. *Assessment, 20*(10), 1-17.

Buchanan, T., & Smith, J. L. (1999). Using the Internet for psychological research: Personality testing on the World Wide Web. *British Journal of Psychology, 90,* 125-144.

Carini, R. M., Hayek, J. C., Kuh, G. D., Kennedy, J. M., & Ouimet, J. A. (2003). College students responses to web and paper surveys: Does mode matter? *Research in Higher Education, 44*(1), 1-19.

Cannon, T. (2012). An introduction to color theory for web designers. Retrieved April 17, 2013 from http://webdesign.tutsplus.com/articles/design-theory/an-introduction-to-color-theory-for-web-designers/.

Carlbring, P., Brunt, S., Bohman, S., Austin, D., Richards, J., Ost, L. –G., & Andersson, G. (2007). Internet vs. paper-and-pencil administration of questionnaires commonly used in panic/agoraphobia research. *Computers in Human Behavior, 23,* 1421-1434.

Cassidy, S., Eachus, P. (2002). Developing the Computer User Self-Efficacy (CUSE) scale: Investigating the relationship between computer self-efficacy, gender and experience with computers. *Journal of Educational Computing Research, 26*(2), 133-153.

Christian, L. M., Dillman, D. A., & Smyth, J. D. (2007). Helping respondents get it right the first time: The influence of words, symbols, and graphics in web surveys. *Public Opinion Quarterly, 71*(1), 113-125.

Cobanoglu, C., Warde, B., and Moreo, P. J. (2001). A comparison of mail, fax, and web survey methods. *International Journal of Market Research, 43*, 405-410.

Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly, 19*(2), 129-211.

Conrad, A. M., & Munro, D. (2008). Relationships between computer self-efficacy, technology, attitudes, and anxiety: Development of the computer technology use scale (CTUS). *Journal of Educational Computing Research, 39*(1), 51-73.

Conrad, F. G., Couper, M. P., & Tourangeau, R. (2003). Interactive features in web surveys. *Joint Meetings of the American Statistical Association,* San Francisco, CA, August 6, 2003.

Cook, C., Heath, F., & Thompson, R. L. (2000). A meta-analysis of response rates in web- or Internet- based surveys. *Educational and Psychological Measurement, 60*, 821-836.

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64*(4), 464-494.

Couper, M. P. (2001). Web surveys: The questionnaire design challenge. *Proceedings of the 53rd Session of the International Statistical Institute, Seoul, Korea*.

Couper, 2011. The future of modes of data collection. *Public Opinion Quarterly, 75*(5), 889-908.

Couper, M. P., & Miller, P. V. (2008). Web survey methods. *Public Opinion Quarterly, 72*(5), 831-835.

Couper, M. P., Tourangeau, R., & Kenyon, K. (2004). Picture this! Exploring visual effects in web surveys. *Public Opinion Quarterly, 68*(2), 255-266.

Couper, M. P., Traugott, M. W., & Lamias, M. J. (2001). Web survey design and administration. *The Public Opinion Quarterly, 65*(2), 230-253.

Curasi, C. F. (2001). A critical exploration of face-to-face interviewing vs. computer-mediated interviewing. *International Journal of Market Research, 43*(4), 361-375.

Daley, E. M., McDermott, R. J., McCormack-Brown, K. R., & Kittleson, M. J. (2003). Conducting web-based survey research: A lesson in Internet designs. *American Journal of Health Behavior, 27*(2), 116-124.

de Beuckelaer, A., & Lievens, F. (2008). Measurement equivalence of paper-and-pencil and Internet organisational surveys: A large-scale examination of 16 countries. *Applied Psychology: An International Review, 58*(2), 336-361.

Denscombe, M. (2006). Web-based questionnaires and the mode effect: An evaluation based on completion rates and data contents of near-identical questionnaires delivered in different modes. *Social Science Computer Review, 24*, 246-254.

Deutskens, E., Ruyter, K. D., Wetzels, M., Oosterveld, P. (2004). Response rate and response quality of Internet-based surveys: An experimental study. *Marketing letters*, 15(1), 21-36.

Dillman, D. A., & Christian, L. M. (2005). Survey mode as a source of instability in response across surveys. *Field Methods, 17*(1), 30-52.

Dillman, D. A., Tortora, R.D., Conradt, J., & Bowker, D. (1998). Influence of plain vs. fancy design on response rates for web surveys. *Proceedings of the Survey Research Methods Section of the American Statistical Association*. Retrieved May 7, 2013 from http://www.sesrc.wsu.edu/dillman/papers/1998/influenceofplain.pdf.

Djamasbi, S., Siegel, M., & Tullis, T. (2010). Generation Y, web design, and eye tracking. *International Journal of Human-Computer Studies, 68*, 307-323.

Fan, W., & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior, 26*, 132-139.

Fowler, F. J., Jr., (2009). Methods of data collection. In *Survey research methods (4th ed.)*, pp. 69-86. Sage Publications: Thousand Oaks, CA.

Fricker, R. D., & Schonlau, M. (2002). Advantages and disadvantages of internet research surveys: Evidence from the literature. *Field Methods, 14*(4), 347-367.

Gaddis, S. E. (1998). How to design online surveys. *Training & Development,* June, 67-72.

Gannon, M. J. (1973). The proper use of the questionnaire survey. *Business Horizons, 16*(5), 89-94.

Garcia, M. L., & Caldera, C. I. (1996). The effect of color and typeface on the readability of on-line text. *Computers and Industrial Engineering, 31*(1/2), 519-524.

Gay, L. R., Mills, G. E., & Airasian, P. W. (2012). Survey research. In *Educational research: Competencies for analysis and applications, 10th ed.* (pp. 183-196). Boston: Pearson.

Godar, S. H. (2000). Use of color and responses to computer-based surveys. *Perceptual and Motor Skills, 91*, 767-770.

Gressard, C. P., & Loyd, B. H. (1985). Validation studies of a new Computer Attitudes Scale. *Paper presented at the [69th] Annual Meeting of the American Educational Research Association, Chicago, IL, March 31-April 4, 1985*.

Groves, R. M., Presser, S., & Dipko, S. (2004). The role of topic interest in survey participation decisions. *The Public Opinion Quarterly*, *68*(1), 2-31.

Gwaltney, C. J., Shields, A. L., & Shiffman, S. (2008). Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: A meta-analytic review. *Value in Health, 11*(2), 322-333.

Hardré, P. L., Crowson, H. M., & Xie, K. (2012). Examining contexts-of-use for web-based and paper-based questionnaires. *Educational and Psychological Measurement, 72*(6), 1015-1038.

Hartley, J., & Rutherford, A. (2003). The effects of using colored paper to boost response rates to surveys and questionnaires. *Journal of Technical Writing and Communication, 33*(1), 29-40.

Heerwegh, D. (2009). Mode differences between face-to-face and web surveys: An experimental investigation of data quality and social desirability effects. *International Journal of Public Opinion Research, 21*(1), 111-121.

Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus web surveying in a high-Internet-coverage population. *Public Opinion Quarterly, 72*(5), 836-846.

Heinssen, R.K., Glass, C.R., & Knight, L.A. (1987). Assessing computer anxiety: development and validation of the computer anxiety rating scale. *Computers and Human Behavior, 3*, 49-59.

Hill, A. L., & Scharff, L. F. (1999). Legibility of computer displays as a function of colour, saturation, and texture backgrounds. *Proceedings for the Second International Conference for Engineering Psychology and Cognitive Ergonomics*.

Howell, R. T., Rodzon, K. S., Kurai, M., & Sanchez, A. H. (2010). A validation of well-being and happiness surveys for administration via the Internet. *Behavior Research Methods, 42*(3), 775-784.

Humar, I., Gradišar, M., & Turk, T. (2008). The impact of color combinations on the legibility of a Web page text presented on CRT displays. *International Journal of Industrial Ergonomics, 38,* 885-899.

Jäckle, A., Roberts, C., Lynn, P. (2010). Assessing the effect of data collection mode on measurement. *International Statistical Review, 78*(1), 3-20.

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

Joinson, A. N., Woodley, A., & Reips, U. –D. (2007). Personalization, authentication, and self-disclosure in self-administered Internet surveys. *Computers in Human Behavior, 23*, 275-285.

Jones, S., Yale, C. J-., Millermaier, S., & Pérez, F. S. (2008). Academic work, the Internet, and U.S. college students. *Internet and Higher Education, 11*, 165-177.

Kaplowitz, M. D., Hadlock, T. D., & Levine, R. (2004). A comparison of web and mail survey response rates. *Public Opinion Quarterly, 68*(1), 94-101.

Karabenick, S. A., et al. (2007). Cognitive processing of self-report items in educational research: Do they think what we mean? *Educational Psychologist, 42*(3), 139-151.

Kay, R. H. (1989). A practical and theoretical approach to assessing computer attitudes: The computer attitude measure (CAM). *Journal of Research on Computing in Education, 21*(4), 456-463.

Kim, D. –H., & Huynh, H. (2008). Computer-based and paper-and-pencil administration mode effects on a statewide end-of-course English test. *Educational and Psychological Measurement, 68*(4), 554-570.

Kinzie, M. B., Delcourt, M. B., & Powers, S. M. (1994). Computer technologies: Attitudes and self-efficacy across undergraduate difficulties. *Research in Higher Education, 35*(6), 745-768.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*, 213-236.

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50*, 536-567.

Kuenzi, J. J. (1999-2000). Trends in college students' computer use and ownership. *Journal of Educational Technology Systems, 28*(1), 21-31.

Kwak, N., & Radler, B. (2002). A comparison between mail and web surveys: Response pattern, respondent profile, and data quality. *Journal of Official Statistics, 18*(2), 257-273.

Kyrnin, J. (2008). Contrasting foreground and background colors: Visualize how the colors display. *Computing: Web design/HTML.* Retrieved April 17, 2013 from http://webdesign.about.com/od/color/l/bl_contrast_table.htm.

Lewis, I., Watson, B., White, K. M. (2009). Internet versus paper-and-pencil survey methods in psychological experiments: Equivalence testing of participant responses to health-related messages. *Australian Journal of Psychology, 61*(2), 107-116.

Lietz, P. (2010). Research into questionnaire design: A summary of the literature. *International Journal of Market Research, 52*(2), 249-272.

Lighthouse International. (2014). Accessibility. Retrieved from http://www.lighthouse.org/accessibility/.

Ling, J., & van Schaik, P. (2002). The effect of text and background color on visual search of Web pages. *Displays, 23*, 223-230.

Ling, J., & van Schaik, P. (2006). The influence of font and line length on visual search and information retrieval in web pages. *International Journal of Human-Computer Studies, 64*, 395-404.

Ling, J., & van Schaik, P. (2007). The influence of line spacing and text alignment on visual search of Web pages. *Displays, 28*, 60-67.

Lugtig, P., Lensvelt-Mulders, G. J., Frerichs, R., & Greven, A. (2011). Estimating non-response bias and mode effects in a mixed-mode survey. *International Journal of Market Research, 53*(5), 669-686.

Lynch, P. J., & Horton, S. (2009). *Web style guide: Basic design principles for creating web sites* (3rd ed.). Available from http://www.webstyleguide.com/.

Malhotra, N. (2008). Completion time and response order effects in web surveys. *Public Opinion Quarterly, 72*(5), 914-934.

McCoy, C. (2010). Perceived self-efficacy and technology proficiency in undergraduate college students. *Computers and Education, 55*, 1614-1617.Medway, R. L., & Fulton, J. (2012). When more gets you less: A meta-analysis of the effect of concurrent web options on mail survey response rates. *Public Opinion Quarterly, 76*(4), 733-746.

Medway, R. L., & Fulton, J. (2012). When more gets you less: A meta-analysis of the effect of concurrent web options on mail survey response rates. *Public Opinion Quarterly, 76*(4), 733-746.

Meier, S. T. (1988). Predicting individual differences in performance on computer-administered tests and tasks: Development of the Computer Aversion Scale. *Computers in Human Behavior, 4*(3), 175-187.

Messineo, M., & DeOllos, I. Y. (2005). Are we assuming too much? Exploring students' perceptions of their computer competence. *College Teaching, 53*(2), 50-55.

Middlestadt, S. E., Bhattacharyya, K., Rosenbaum, J., Fishbein, M., & Shepherd, M. (1974). The use of theory based semistructured elicitation questionnaires: Formative research for CDS's prevention marketing initiative. *Public Health Reports, 111*, 18-27.

Millar, M. M., & Dillman, D. A. (2011). Improving response to web and mixed-mode surveys. *Public Opinion Quarterly, 75*(2), 249-269.

Miller, T. W. (2001). Can we trust the data of online research? *Marketing Research, 13*(2), 26-32.

Mitra, A., Jalin-Shukla, P., Robbins, A., Champion, H., & DuRant, R. (2008). Differences in rate of response to web-based surveys among college students. *International Journal on E-Learning, 7*(2), 265-281.

Moos, D. C., Azevedo, R. (2009). Learning with computer-based learning environments: A literature review of computer self-efficacy. *Review of Educational Research, 79*(2), 576-600.

Morgan, G. A., & Harmon, R. J. (2001). Data collection techniques. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(8), 973-976.

Murphy, C. A., Coover, D., & Owen, S. V. (1989). Development and validation of the computer self-efficacy scale. *Educational and Psychological Measurement, 49*, 893-899.

National Science Foundation (1997). Analyzing qualitative data. Retrieved from http://www.nsf.gov/pubs/1997/nsf97153/chap_4.htm.

O'Rourke, T. (2011). Looking back but moving forward: The opportunities and challenges of a mixed mode approach to survey research. *American Journal of Health Studies, 26*(2), 114-118.

Oliver, K. (2002). Web interface design. Retrieved March 20, 2013 from http://www.edtech.vt.edu/id/interface/index.html.

Oklahoma State University (OSU) Institutional Research and Information Management (IRIM). (2013). Retrieved from http://irim.okstate.edu/images/StudentProfile/2013/pdf/2013psb.pdf.

Oklahoma State University (OSU) System. (2013). Mission statement. Retrieved from http://system.okstate.edu.

Oppermann, M. (1995). E-mail surveys: Potentials and pitfalls. *Marketing Research, 7*(3), 28-33.

Papanastasiou, E. C. (2005). Factor structure of the "Attitudes Toward Research" scale. *Statistics Education Research Journal, 4*(1), 16-26.

Passmore, C., Dobbie, A. E., Parchman, M., & Tysinger, J. (2002). Guidelines for constructing a survey. *Family Medicine, 34*(4), 281-286.

Patton, (2002).

Perkins, G. H. (2004). Will libraries' web-based survey methods replace existing non-electronic survey methods? *Information Technology and Libraries, 23*(3), 123-126.

Pouwer, F., Snoek, F. J., van der Ploeg, H. M., Heine, R. J., & Brand, A. N. (1998). A comparison of the standard and the computerized versions of the Well-being Questionnaire (WBQ) and the Diabetes Treatment Satisfaction Questionnaire (DTSQ). *Quality of Life, 7*, 33-38.

Public Opinion Quarterly. (1973-present). About this journal. Retrieved on April 17, 2013 from http://poq.oxfordjournals.org/.

Prunk, T. (1994). The value of questionnaires. *Target Marketing, 17*(10), 37-38.

Qualitative data analysis. [Lecture]. Retrieved on March 6, 2014 from http://www.southalabama.edu/coe/bset/johnson/lectures/lec17.pdf.

Qualtrics. [Computer software]. (2014). Available from https://qualtrics.com/

Redline, C. D., & Dillman, D. A. (1999). *The influence of auxiliary, symbolic, numeric, and verbal languages on navigational compliance in self-administered questionnaires*. Paper presented at the International Conference on Survey Nonresponse, University of Lausanne, Switzerland.

Redline, C.D., Dillman, D. A., Carley-Baxter, L., & Creecy, R. (2003). Factors that influence reading and comprehension of branching instructions in self–administered questionnaires. *Advances in Statistical Analysis, 89*(1), 21-38

Richman, W. L., Weisband, S., Kiesler, S., Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology, 84*(5), 754-775.

Roster, C. A., Rogers, R. D., Hozier, G. C., Baker, K. G., & Albaum, G. (2007). Management of marketing research projects: Does delivery mode matter anymore in survey research? *Journal of Marketing Theory and Practice, 15*(2), 127-144.

Ryan, J. M., Corry, J. R., Attewell, R., & Smithson, M. J. (2002). A comparison of an electronic version of the SF-36 General Health Questionnaire to the standard paper version. *Quality of Life, 11,* 19-26.

Sax, L. J., Gilmartin, S. K., & Bryant, A. N. (2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education, 44*(4), 409-432.

Schaeffer, N. C., & Dykema, J. (2011). Questions for surveys: Current trends and future directions. *Public Opinion Quarterly, 75*(5), 909-961.

Schmidt, K. E., Liu, Y., & Sridharan, S. (2009). Webpage aesthetics, performance, and usability: Design variables and their effects. *Ergonomics, 52*(6), 631-643.

Schmidt, W. C. (1997). World-wide web survey research: Benefits, potential problems, and solutions. *Behavior Research Methods, 29*(2), 274-279.

Schutt, R.K. (2012). Qualitative data analysis. In *Investigating the social world* (Chapter 10). Retrieved March 7, 2014 from http://www.sagepub.com/schuttisw7e/study/intro.htm.

Schwarz, N., Strack, F., Hippler, H. –J., & Bishop, G. (1991). The impact of administration mode on response effects in survey measurement. *Applied Cognitive Psychology, 5*, 193-212.

Section508.gov. (1998). Section 508 of the Rehabilitation Act of 1973 (29 U.S.C. § 794d). Retrieved March 29, 2014 from http://www.section508.gov/Section-508-Of-The-Rehabilitation-Act.

Sethuraman, R., Kerin, R. A., & Cron, W. L. (2005). A field study comparing online and offline data collection methods for identifying product attribute preferences using conjoint analysis. *Journal of Business Research, 58*, 602-610.

Shapiro, J. S., Bessette, M. J., Baumlin, K. M., Ragin, D. F., & Richardson, L. D. (2004). Automating data collection. *Academic Emergency Medicine, 11*(11), 1223-1228.

Shropshire, K. O., Hawdon, J. E., & Witte, J. C. (2009). Web survey design: Balancing measurement, response, and topical interest. *Sociological Methods & Research, 37*, 344-370.

Singh, A., Taneja, A., & Mangalaraj, G. (2009). Creating online surveys: Some wisdom from the Trenches tutorial. *IEEE Transactions on Professional Communication, 52*(2), 197-212.

Stanton, J. M. (1998). An empirical assessment of data collection using the Internet. *Personnel Psychology, 51*, 709-725.

Stern, M. J., Smyth, J. D., & Mendez, J. (2012). The effects of item saliency and question design on measurement error in a self-administered survey. *Field Methods, 24*(1), 3-27.

Sue, V. M., & Ritter, L. A. (2007). Conducting online surveys (2nd ed.). Thousand Oaks, CA: Sage Publications.

Torkzadeh, G., Chang, J. C-J., & Demirhan, D. (2006). A contingency model of computer and Internet self-efficacy. *Information & Management, 43*, 541-550.

Torkzadeh, G., & van Dyke, T. P. (2002). Effects of training on Internet self-efficacy and computer use attitudes. *Computers in Human Behavior, 18*, 479-494.

Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology, 55*, 775-801.

Vallejo, M. A., Mañanes, G., Comeche, M. I., & Diaz, M. I. (2008). Comparison between administration via Internet and paper-and-pencil administration of two clinical instruments: SCL-90-R and GHQ-28. *Journal of Behavior Therapy, 39*, 201-208.

Vannieuwenhuyze, J., & Loosveldt, G. (2012). Evaluating relative mode effects in mixed-mode surveys: Three methods to disentangle selection and measurement effects. *Sociological Methods & Research, 42*(1), 82-104.

Vannieuwenhuyze, J., Loosveldt, G., & Molenberghs, G. (2010). A method for evaluating mode effects in mixed-mode surveys. *Public Opinion Quarterly, 74*(5), 1027-1045.

Vereecken, C. A., & Maes, L. (2006). Comparison of a computer-administered and paper-and-pencil administered on health and lifestyle behaviors. *Journal of Adolescent Health, 38*, 426-432.

Vicente, P., & Reis, E. (2010). Using questionnaire design to fight nonresponse bias in web surveys. *Social Science Computer Review, 28*(2), 251-267.

Walston, J. T., Lissitz, R. W., & Rudner, L. M. (2006). The influence of web-based questionnaire presentation variation on survey cooperation and perceptions of survey quality. *Journal of Official Statistics, 22*(2), 271-291.

Ward, P., Clark, T., Zabriskie, R., & Morris, T. (2012). Paper/pencil versus online data collection: An exploratory study. *Journal of Leisure Research, 44*(4), 507-530.

Ware, C. (2000). Information visualization: Perception for design. San Francisco: Morgan Kaufmann.

WebAIM. (2014a) Color contrast checker. Retrieved on March 6, 2014 from http://webaim.org/resources/contrastchecker/.

WebAIM. (2014b) Introduction to accessibility. Retrieved on March 6, 2014 from http://webaim.org/intro/.

WebAIM. (2014c). United States laws. Retrieved from http://webaim.org/articles/laws/usa/rehab.

Webb, J. (2000). Questionnaires and their design. *The Marketing Review, 1*, 197-218.

Webster, J., & Compeau, D. (1996). Computer-assisted versus paper-and-pencil administration of questionnaires. *Behavior Research Methods, Instruments, and Computers, 28*(4), 567-576.

Weible, R., & Wallace, J. (1998). Cyber research: The impact of the Internet on data collection. *Marketing Research, 10*(3), 19-31.

Weigold, A., Weigold, I. K., & Russell, E. J. (2013). Examination of the equivalence of self-report survey-based paper-and-pencil and Internet data collection methods. *Psychological Methods, 18*(1), 53-70.

Weston, T. J., Barker, L. J. (2002). A profile of student computer use, training, and proficiency. *Journal of Computing in Higher Education, 14*(1), 87-112.

Wood, E., Nosko, A., Desmarais, S., Ross, C., & Irvine, C. (2006). Online and traditional paper-and-pencil survey administration: Examining experimental presence, sensitive material, and long surveys. *The Canadian Journal of Human Sexuality, 15*(3/4), 147-155.

Wright, K. B. (2005). Researching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication, 10*(3).

APPENDIX A

Government Knowledge Test

What is freedom of religion?

- You must choose a religion.
- You can practice any religion, or not practice a religion.
- You can't choose the time you practice your religion.
- No one can practice a religion.

We elect a President for how many years?

- Four (4)
- Two (2)
- Ten (10)
- Eight (8)

What do we show loyalty to when we say the Pledge of Allegiance?

- The state where you live
- The President
- Congress
- The United States

Who lived in America before the Europeans arrived?

- American Indians
- Floridians
- No one
- Canadians

Name the U.S. war between the North and the South.

- World War I
- The Revolutionary War
- The Civil War
- The War of 1812

Why does the flag have 13 stripes?

Because the stripes represent the members of the Second Continental Congress
Because it was considered lucky to have 13 stripes on the flag
Because the stripes represent the original colonies
Because the stripes represent the number of signatures on the U.S. Constitution

What do we call the first ten amendments to the Constitution?

The Declaration of Independence
The Bill of Rights
The inalienable rights
The Articles of Confederation

What is an amendment?

The beginning of the Declaration of Independence
The Preamble to the Constitution
An addition (to the Constitution)
An introduction

Name one branch or part of the government.

United Nations
Parliament
State government
Legislative

What are the two parts of the U.S. Congress?

The Senate and House of Representatives
The House of Representatives and the courts
The House of Lords and the House of Commons
The Senate and the courts

What are two rights of everyone living in the United States?

Freedom to petition the government and freedom to disobey traffic laws
Freedom of worship and freedom to make treaties with other countries
Freedom of speech and freedom of worship
Freedom of speech and freedom to run for president

What group of people was taken to America and sold as slaves?

English
Canadians
Dutch
Africans

When was the Declaration of Independence adopted?

December 7, 1787
July 4, 1789
July 4, 1776
March 4, 1789

There were 13 original states. Name three.

New York, Kentucky, and Georgia
Washington, Oregon, and California
Maryland, Virginia, and North Carolina
Virginia, North Carolina, and Florida

Who is the "Father of Our Country"?

Abraham Lincoln
Thomas Jefferson
Patrick Henry
George Washington

What did the Emancipation Proclamation do?

Ended World War I
Freed slaves in most Southern states
Gave women the right to vote
Gave the United States independence from Great Britain

Name one American Indian tribe in the United States.

Celts
Cherokee
Zawi Chemi
Slavs

Name one of the two longest rivers in the United States.

Ohio River
Rio Grande River
Colorado River
Mississippi River

What is the name of the national anthem?

God Bless the U.S.A.
The Star-Spangled Banner
My Country Tis of Thee
America the Beautiful

Name two national U.S. holidays.

Valentine's Day and Presidents' Day
April Fool's Day and Labor Day
Labor Day and Thanksgiving
Citizenship Day and Columbus Day

APPENDIX B

STUDENTS' "ATTITUDES TOWARD RESEARCH" SCALE

The following statements refer to some aspects of educational research. Please answer all the questions sincerely. **DO NOT DISCLOSE YOUR IDENTITY ANYWHERE**.

Circle one of the numbers opposite each of the statements that follow.
(By selecting number 1, you indicate that you strongly agree).
(By selecting number 7, you indicate that you strongly disagree).

Statement		Strongly agree						Strongly disagree
Research makes me anxious.		1	2	3	4	5	6	7
Research should be taught to all students.		1	2	3	4	5	6	7
I enjoy research.		1	2	3	4	5	6	7
Research is interesting.		1	2	3	4	5	6	7
I like research.		1	2	3	4	5	6	7
I feel insecure concerning the analysis of research data.		1	2	3	4	5	6	7
Research scares me.		1	2	3	4	5	6	7
Research is useful for my career.		1	2	3	4	5	6	7
I find it difficult to understand the concepts of research.		1	2	3	4	5	6	7
I make many mistakes in research.		1	2	3	4	5	6	7
I have trouble with arithmetic.		1	2	3	4	5	6	7
I love research.		1	2	3	4	5	6	7
I am interested in research.		1	2	3	4	5	6	7
Research is connected with my field of study.		1	2	3	4	5	6	7

Most students benefit from research.		1	2	3	4	5	6	7
Research is stressful.		1	2	3	4	5	6	7
Research is very valuable.		1	2	3	4	5	6	7
Research makes me nervous.		1	2	3	4	5	6	7
I use research in my daily life.		1	2	3	4	5	6	7
The skills I have acquired in research will be helpful to me in the future.		1	2	3	4	5	6	7
Research is useful to every professional.		1	2	3	4	5	6	7
Knowledge from research is as useful as writing.		1	2	3	4	5	6	7
Research is irrelevant to my life.		1	2	3	4	5	6	7
Research should be indispensable in my professional training.		1	2	3	4	5	6	7
Research is complicated.		1	2	3	4	5	6	7
Research thinking does not apply to my personal life.		1	2	3	4	5	6	7
I will employ research approaches in my profession.		1	2	3	4	5	6	7
Research is difficult.		1	2	3	4	5	6	7
I am inclined to study the details of research procedures carefully.		1	2	3	4	5	6	7
Research is pleasant.		1	2	3	4	5	6	7
Research-oriented thinking plays an important role in my daily life.		1	2	3	4	5	6	7
Research is a complex subject.		1	2	3	4	5	6	7

APPENDIX C

COMPUTER SELF-EFFICACY SCALE

The purpose of this questionnaire is to examine attitudes toward the use of computers. The questionnaire is divided into two parts. In Part 1, you are asked to provide some basic background information about yourself and your experience of computers, if any. Part 2 aims to elicit more detailed information by asking you to indicate the extent to which you, personally, agree or disagree with the statements provided.

Part 1:

Age:

○	○	○	○	○	○
18-24	25-31	32-38	39-45	46-52	53+

Experience with computers

○	○	○	○	○	○
None	Very limited	Some experience	Quite a lot	Extensive	53+

Please indicate the computer packages (software) you have used. Check all that apply.

○	○	○	○	○	○	○	○
Word processing packages	Spreadsheets	Databases	Presentation packages	Statistics packages	Desktop publishing	Multimedia	Other, please specify

Do you own a computer?

○	○
Yes	No

Have you ever attended a computer training course?

○	○
Yes	No

Part 2:

Below you will find a number of statements concerning how you might feel about computers. Please indicate the strength of your agreement/disagreement with the statements using the 6-point scale shown below. Tick the box (i.e., between 1 and 6) that most closely represents how much you agree or disagree with the statement. There are no *correct* responses; it is your own views that are important.

Statement		Strongly disagree					Strongly agree
Most difficulties I encounter when using computers, I can usually deal with.		1	2	3	4	5	6
I find working with computers very easy.		1	2	3	4	5	6
I am very unsure of my abilities to use computers.		1	2	3	4	5	6
I seem to have difficulties with most of the packages I have tried to use.		1	2	3	4	5	6
Computers frighten me.		1	2	3	4	5	6
I enjoy working with computers.		1	2	3	4	5	6
I find that computers get in the way of learning.		1	2	3	4	5	6
DOS-based computer packages don't cause many problems for me.		1	2	3	4	5	6
Computers make me much more productive.		1	2	3	4	5	6
I often have difficulties when trying to learn how to use a new computer package.		1	2	3	4	5	6
Most of the computer packages I have had experience with have been easy to use.		1	2	3	4	5	6
I am very confident in my abilities to make use of computers.		1	2	3	4	5	6
I find it difficult to get computers to do what I want them to do.		1	2	3	4	5	6
At times I find working with computers very confusing.		1	2	3	4	5	6
I would rather that we did not have to learn how to use computers.		1	2	3	4	5	6
I usually find it easy to learn how to use a new software package.		1	2	3	4	5	6
I seem to waste a lot of time struggling with computers.		1	2	3	4	5	6

Using computers makes learning more interesting.		1	2	3	4	5	6
I always seem to have problems when trying to use computers.		1	2	3	4	5	6
Some computer packages definitely make learning easier.		1	2	3	4	5	6
Computer jargon baffles me.		1	2	3	4	5	6
Computers are far too complicated for me.		1	2	3	4	5	6
Using computers is something I rarely enjoy.		1	2	3	4	5	6
Computers are good aids to learning.		1	2	3	4	5	6
Sometimes, when using a computer, things seem to happen and I don't know why.		1	2	3	4	5	6
As far as computers go, I don't consider myself to be very competent.		1	2	3	4	5	6
Computers help me to save a lot of time.		1	2	3	4	5	6
I find working with computers very frustrating.		1	2	3	4	5	6
I consider myself to be a skilled computer user.		1	2	3	4	5	6
When using computers, I worry that I might press the wrong button and damage it.		1	2	3	4	5	6

.

APPENDIX D

Demographic Information Sheet

Providing the following information is optional. Thank you for your participation.

Personal Information

College classification:

○	○	○	○	○
Freshman	Sophomore	Junior	Senior	Graduate student

College major:

Computer usage

Please estimate the amount of time you spend on the computer for professional or educational purposes (include time spent on Internet):

○	○	○	○	○
>1 hour	1 to 3 hours	4 to 6 hours	7 to 9 hours	9+ hours

Please indicate the type of device you are currently using to complete this survey:

○	○	○	○	○
Personal computer (desktop)	Personal computer (laptop)	Tablet	Mobile phone	Other

Other information

Are you currently wearing contacts or glasses?

○	○
Yes	No

Please provide any comments you may have regarding the survey process, including your thoughts on the visual design elements (e.g., color of font, color of background, use of the progress bar indicator, and use of the highlight bar) of the survey.

Thank you for taking the time to fill out this portion of the survey.

APPENDIX E

Participants' feedback to the following question:

"*Please provide any comments you may have regarding the survey process, including your thoughts on the visual design elements (e.g., color of font, color of background, use of the progress bar indicator, and use of the highlight bar) of the survey*."

Visual design	**Participant feedback**
Red font; black background; highlight bar present; progress bar present	Nice shade of yellow, although the colors started to annoy me (might be PMS though). When a question is answered and while trying to move to the next question using down arrow button, the selected option is being modified which should not be the case for easy navigation. Also the colors need to be simpler. Just black and white would be better than red and yellow. Thank You. The dark background really keeps my eyes from straining as much and the red font is appealing. The yellow is a little too bright for me. On the multiple answer questions about programs it was extremely hard to see the white text with the yellow background. I like the way it looks overall though! Survey was very hard to read! This survey was not enjoyable at all, same question in different iterations over and over again, what kind of subliminal instrument is this, it just makes respondents go through the survey faster as to hurry it along.
Red font; black background; highlight bar absent; progress bar present	This red font is very distracting. It made me have to concentrate harder on the questions. It also made me antsy and nervous. Don't ever do this to people again. I didn't like the red font of the survey because it messed with my eyes for a little bit. The red font is distracting and harder to read. I did not care for the color. The black and red color scheme made the overall presentation very serious, giving it a sense of professionalism; it was a very enjoyable experience taking this survey.
Blue font;	Do not like the highlight bar, the blue on black text is difficult to read.

black background; highlight bar present; progress bar present	It was hard to see with the black background and blue font. The highlighting helped but it was too bright. I would think that simple colors would be ideal in research circumstances. The visual appearance is horrible and the lay out is not appealing I understand that you were trying to help improve focus on one question at a time with your dark background, yellow highlighting, and blue text. This provides the desired result. However, I would encourage a less sharp contrast between the in focus and out of focus question. Given the fast movements of our eyes, we don't look at one thing at once and the color scheme is fighting that tendency. The increased entropy makes it harder to focus on the one question we can easily see because our mind and eyes are trying to also read the things that are not highlighted. It is the skimming nature of reading. That is why books can have lots of sentences on a single page. I did not like the color of the font or the color the background. I found it distracting. Couldn't stand the blue on black type, thought about quitting survey. The color of the survey is very ugly and unattractive; it has been bothering me throughout the survey. I hated the blue colored font and the black background, it wasn't interesting at all. And the questions are too much. The questions about USA i answered all of them googling and it made me mad doing that because it took a lot of my time. The yellow was extremely bright against the black background.
Blue font; black background; highlight bar absent; progress bar present	It was extremely difficult to read. The blue text was too dark and the black background. Distraction. I haven't seen blue font used on surveys much. The progress bar indicator is rather small and not very important it seems. The background looks like a blurred image. The Qualtrics 'Q' is cut-off at the top which bugs me (OCD). The rest of the survey seemed fine, just a few differences that i did not notice too much until further inspection. The black makes the blue pop, but it makes it difficult to read in comparison to the usual white/black text. Background interesting, completion bar was mildly informative, first section was a little repetitive.

	My friend is colorblind and he says he would rather kill himself than use this color scheme. Also, The darker part of the spectrum was much more pleasing than the lighter background with the blue text, whereas the lighter background was more offensive and irritating to look at, and probably could have an effect on my choices for answers. I did try to be consistent with inverse questions, given that I recognized the alternate wording and answered the same given the same wording. Also, there were some similar questions that instead of similar wording, a key word or phrase was replaced that changed the meaning of the questions, in which case I answered differently, or to a different degree, because of the wording. Also, some questions, for instance *computers sometimes do things that I don't understand*(paraphrased), I was forced to answer strongly agree, because, especially when I'm writing code, it just sometimes works, and God knows why! Regardless, I hope this has helped you in your research endeavors. The blue font color clashes with the gray background.
Red font; black background; highlight bar present; progress bar absent	Yellow may be a bit bright on mobile phone, but there wasn't much difficulty reading the red text over it. The colors are interesting. The red on black was cool, but the yellow hurts my eyes. At one point there was white on yellow and that was nearly impossible to read. The Yellow Indication bar is a bit violent and abrasive. The use of bright red font with a bright yellow highlight is painful to the eyes and makes the text significantly less legible. The highlighted area caught me off-guard at first, but then I liked it because it helped me focus on the one question. The red text on the black screen was difficult to read - it made the font seem even smaller. Harder to see with the yellow background. Good highlighting.
Red font; black background; highlight bar absent; progress bar absent	Red just sucks to use when you are on the computer. I know lots of people who have trouble seeing red. I don't and this survey was annoying. Red font kept my attention on the text rather than there being a distracting background. This red font is TERRIBLE!!!! Also the changing black to grey to black background is frustrating, but I assume it has a underlying

	purpose. Also Junior is spelled wrong. The Background changing colors from dark to light distracted and confused me. Missed the progress bar, don't like the red on black, wavy background in black makes it harder to see. Red font not best choice. What progress bar? The black and red are awful color choices for a survey, almost made me want to just exit. My eyes don't like the contrast. The red and black chosen in this survey was difficult to read at times. Red on Black is rather difficult to read. At least it's not gray on gray.
Blue font; black background; highlight bar present; progress bar absent	I [could] barely see the question in the blue font on the black background. Very annoying colors used. Without the highlight bar, this survey would have been nigh impossible to complete, due to illegibility. Even with it, a combination of blue text on a yellow background made it somewhat uncomfortable to view the questions and answers. Color of font hurts my eyes, dark background is better than white, didn't see a progress bar indicator, the highlight bar is stressful and awful. In situations where it was bright outside i had difficulty reading what i was about to select and only after clicking was i able to fully read the question and answers. I went into the survey with a slight headache and now it is far worse. Difficult to read.
Blue font; black background; highlight bar absent; progress bar absent	Although difficult to read at times the color was light on the eyes and patterns in the questioning allowed for quick responses. A little hard to see. The font was extremely difficult to read and it almost made me [exit] out of the survey. I found the blue text off-putting and unprofessional. It was harder to read in some places, and was ugly. I would never use this scheme

	The font [color] and background look like they fell out of the nineties. Awful. Way too hard to read the font on the background provided. I had to tilt my laptop screen just right to be able to read the statements. I don't like the blue font with the background. The royal blue text against the gray made it very difficult to complete. I did not much care when the background lighting was very bright with the dark blue text. It is hard to read, the colors need to be opposites on the color wheel. I found it difficult to read the questions. The background had shades of black; I could read the questions on black background with ease when compared with other shades. And it was basically the angle at which I see the screen that is causing the problem. The blue font on the black background hurt my eyes! The survey questions were good. Visual design elements definitely need to be changed--they were very distracting.
Red font; white background; highlight bar present; progress bar present	The color scheme is garish and annoying. The highlighting is ok, as well as the status bar. The survey completion element on the bottom does not seem to work. Red on white is not a very pleasant color scheme. Black on white, or a deep gray on white would be more preferable. The red text was not enjoyable. Also, the yellow highlights were not fun. Furthermore, the fonts used made me feel uncomfortable. Also, there were too many questions per page. I do not care for the red font, it makes it difficult to read. The yellow that flashes when you click on an answer is distracting and hurts my eyes. I worry that someone who is visually impaired would have difficulty taking this survey. Someone with red green colorblindness would have trouble with this survey. This box is too small to write a long answer in because you cannot see what you have previously typed. The font type is not the easiest to read, but it is a standard font that many use. The progress bar is definitely a good feedback device, as is the highlight bar on the current question. Speaking personally, white is an excessively bright color to have on a computer screen due to eye strain after extended viewing.

	Red color in the letters is Ok, but highlighting them with bright yellow is almost painful to the eyes. Color of font and highlight annoyed me greatly. Gave me a headache.
Red font; white background; highlight bar absent; progress bar present	Font Color sucks… Good. Some people that might have difficulties with computers might not be able to take this... The red font is difficult to read and distracting. The neutral background is fine (but doesn't provide enough contrast with the red font). It would be nice if the progress indicated what numerical percentage of the survey had been completed (rather than just shading the progress bar). Some of the statements (in the computer section) are worded in a somewhat confusing manner (e.g., using extremes - definitely, always). Unrelated, I shudder to think how poorly students will perform on the history section. :(
Blue font; white background; highlight bar present; progress bar present	The font color didn't seem to have any effect on me but the highlight bar made it easier for me to go faster through the survey. The progress bar indicator was helpful to know how far through, and seemed to give me a feeling of accomplishment as I filled out the survey. The blue text on grey background seems more painful for me to read than black would be. The grey border around the questions with a larger darker grey background seemed to help guide me down the page. The yellow highlighting was very bright and caused some squinting but made it easier to see what question I was looking at. It was pretty bland. The yellow is too [aversive]. I like it, help to know where you are in the survey. The color of the font and the highlight bar are very distracting. There is way too much contrast and brightness with the highlight bar, and the blue of the font bothers my eyes because they have become accustomed to reading black font. The progress bar was helpful and the grey background worked well and was not distracting. The progress bar concerned me because it didn't have a percentage complete shown, and I thought the yellow background behind each individual question helped me to see each set of answers and not get lost in all of the questions.

	Seemed pretty simple and straight forward on the questions.
Blue font; white background; highlight bar absent; progress bar present	Color of font is difficult to read. Why not make it black? The survey seemed to operate as intended. The survey became more interesting once the banner was showing. That's a [good] survey. Boring but concise, blue font with grey background. Blue is hard to read. I really liked the font and font color. It was better than the normal dull black.
Red font; white background; highlight bar present; progress bar absent	I'm not a fan of the red font color; the yellow background originally annoyed me. The yellow background no longer annoys me. The bright yellow is jarring. I don't care for the red font or all of the gray space on either side of the questions which takes up too much space. I don't recall seeing a progress bar indicator. Very visible means of writing, and my only negative observation was that unless you clicked on either the text or the actual button, not anywhere along the horizontal line inside the highlighted yellow box, then it wouldn't select the answer, although it would highlight it.
Red font; white background; highlight bar absent; progress bar absent	Short pointed questions that are easy to answer is good - issuing a survey via online survey about computers- may get biased results. Red font is somewhat harder to read. I don't like to read red type. Not a fan of red font.
Blue font; white background; highlight bar present; progress bar absent	Blue font is easier to read than black font and is visually more appealing than black font. I noticed quite a bit of repetition in questions... Highlighting bar helped me keep track of where I was in the survey. Although the white was a bit glaring for my glasses, the blue font was easy to read. Word color with background made the words easy to read and the highlighting after you answered made it easier to transition better to the next question and not lose your place. The background was really bright.

	Redundant questions in the section about research. The yellow bar was obnoxious. Highlight bar was a bit bright. Gray and light gray background nice. Blue text also a bit bright, black would've been sufficient.
Blue font; white background; highlight bar absent; progress bar absent	I like the background color; it is less harsh than a stark white. I don't really like the blue color of the type. The size of the type was fine. Liked the blue it helps my eyes, and the computer 'problem' was funny on the first part. Blue and Grey are bland background colors so no distractions, don’t like the arrow does not go with layout. Use of a progress bar would have been more helpful to gauge time to [complete];the font color was tiring for my eyes. It was really long and some of the questions repeated itself, good luck! Asking "DOS-based" on a survey about difficulties with computers seems odd. I answered, but have no clue what DOS means!!

APPENDIX F

Means and standard deviations for the Government Knowledge test

Font color	Progress bar	Highlight bar	Background color	Mean	Standard deviation	N
Red	Present	Present	Black	19.00	1.41	7
			White	19.21	0.97	14
			Total	19.14	1.11	21
		Absent	Black	19.78	0.44	9
			White	18.70	3.09	10
			Total	19.21	2.27	19
		Total	Black	19.44	1.03	16
			White	19.00	2.09	24
			Total	19.18	1.74	40
	Absent	Present	Black	19.45	0.93	11
			White	19.57	0.53	7
			Total	19.50	0.79	18
		Absent	Black	19.33	0.91	21
			White	19.15	1.63	13
			Total	19.26	1.21	34
		Total	Black	19.38	0.91	32
			White	19.30	1.34	20
			Total	19.35	1.08	52
	Total	Present	Black	19.28	1.13	18
			White	19.33	0.86	21
			Total	19.31	0.98	39
		Absent	Black	19.47	0.82	30
			White	18.96	2.33	23
			Total	19.25	1.65	53
		Total	Black	19.40	0.94	48
			White	19.14	1.77	44
			Total	19.27	1.40	92
Blue	Present	Present	Black	19.18	0.75	11
			White	19.31	0.95	13
			Total	19.25	0.85	24
		Absent	Black	18.93	1.54	14
			White	17.50	3.83	16
			Total	18.17	3.03	30
		Total	Black	19.04	1.24	25
			White	18.31	3.01	29
			Total	18.65	2.37	54
	Absent	Present	Black	18.38	2.78	16
			White	19.00	1.73	19
			Total	18.71	2.26	35
		Absent	Black	18.72	2.08	18

			White	19.35	1.11	17
			Total	19.03	1.69	35
		Total	Black	18.56	2.40	34
			White	19.17	1.46	36
			Total	18.87	1.98	70
	Total	Present	Black	18.70	2.20	27
			White	19.13	1.45	32
			Total	18.93	1.83	59
		Absent	Black	18.81	1.84	32
			White	18.45	2.89	33
			Total	18.63	2.42	65
		Total	Black	18.76	1.99	59
			White	18.78	2.31	65
			Total	18.77	2.16	124
Total	Present	Present	Black	19.11	1.02	18
			White	19.26	0.94	27
			Total	19.20	0.97	45
		Absent	Black	19.26	1.29	23
			White	17.96	3.55	26
			Total	18.57	2.78	49
		Total	Black	19.20	1.17	41
			White	18.62	2.63	53
			Total	18.87	2.13	94
	Absent	Present	Black	18.81	2.25	27
			White	19.15	1.52	26
			Total	18.98	1.92	53
		Absent	Black	19.05	1.57	39
			White	19.27	1.34	30
			Total	19.14	1.47	69
		Total	Black	18.95	1.87	66
			White	19.21	1.41	56
			Total	19.07	1.67	122
	Total	Present	Black	18.93	1.85	45
			White	19.21	1.25	53
			Total	19.08	1.55	98
		Absent	Black	19.13	1.47	62
			White	18.66	2.66	56
			Total	18.91	2.12	118
		Total	Black	19.05	1.63	107
			White	18.93	2.11	109
			Total	18.99	1.88	216

APPENDIX G

Means and standard deviations for 'Attitudes toward Research' scale

Font color	Progress bar	Highlight bar	Background color	Mean	Standard deviation	N
Red	Present	Present	Black	159.43	11.37	7
			White	159.29	24.10	14
			Total	159.33	20.40	21
		Absent	Black	162.89	21.11	9
			White	148.30	22.28	10
			Total	155.21	22.41	19
		Total	Black	161.38	17.10	16
			White	154.71	23.52	24
			Total	157.38	21.20	40
	Absent	Present	Black	166.45	25.37	11
			White	145.29	23.62	7
			Total	159.78	25.49	18
		Absent	Black	160.67	26.22	21
			White	160.77	28.12	13
			Total	160.71	26.54	34
		Total	Black	162.66	25.67	32
			White	156.75	26.59	20
			Total	160.38	25.93	52
	Total	Present	Black	163.72	20.90	18
			White	155.95	23.84	21
			Total	159.54	22.58	39
		Absent	Black	161.33	24.46	30
			White	155.35	25.97	23
			Total	158.74	25.06	53
		Total	Black	162.23	22.99	48
			White	155.64	24.69	44
			Total	159.08	23.91	92
Blue	Present	Present	Black	174.91	22.57	11
			White	160.69	30.06	13
			Total	167.21	27.30	24
		Absent	Black	151.14	30.00	14
			White	152.25	24.14	16
			Total	151.73	26.55	30
		Total	Black	161.60	29.06	25
			White	156.03	26.79	29
			Total	158.61	27.74	54
	Absent	Present	Black	160.50	27.77	16
			White	158.63	23.80	19
			Total	159.49	25.31	35
		Absent	Black	164.44	23.42	18
			White	153.65	26.05	17

			Total	159.20	24.97	35
		Total	Black	162.59	25.24	34
			White	156.28	24.65	36
			Total	159.34	24.96	70
	Total	Present	Black	166.37	26.32	27
			White	159.47	26.07	32
			Total	162.63	26.20	59
		Absent	Black	158.63	26.89	32
			White	152.97	24.76	33
			Total	155.75	25.78	65
		Total	Black	162.17	26.69	59
			White	156.17	25.42	65
			Total	159.02	26.10	124
Total	Present	Present	Black	168.89	20.14	18
			White	159.96	26.61	27
			Total	163.53	24.39	45
		Absent	Black	155.74	26.98	23
			White	150.73	23.07	26
			Total	153.08	24.84	49
		Total	Black	161.51	24.83	41
			White	155.43	25.13	53
			Total	158.09	25.05	94
	Absent	Present	Black	162.93	26.48	27
			White	156.12	23.65	26
			Total	159.58	25.13	53
		Absent	Black	162.41	24.72	39
			White	156.73	26.73	30
			Total	159.94	25.58	69
		Total	Black	162.62	25.25	66
			White	156.45	25.12	56
			Total	159.79	25.28	122
	Total	Present	Black	165.31	24.08	45
			White	157.08	25.04	53
			Total	161.40	24.74	98
		Absent	Black	159.94	25.57	62
			White	153.95	25.06	56
			Total	157.09	25.40	118
		Total	Black	162.20	24.98	107
			White	155.95	25.02	109
			Total	159.05	25.13	216

APPENDIX H

Means and standard deviations for 'Computer User Self-Efficacy' scale

Font color	Progress bar	Highlight bar	Background color	Mean	Standard deviation	N
Red	Present	Present	Black	130.57	26.04	7
			White	148.31	20.70	13
			Total	142.10	23.66	20
		Absent	Black	134.89	22.00	9
			White	126.50	23.85	10
			Total	130.47	22.76	19
		Total	Black	133.00	23.11	16
			White	138.83	24.26	23
			Total	136.44	23.67	39
	Absent	Present	Black	141.55	21.39	11
			White	151.71	20.64	7
			Total	145.50	21.11	18
		Absent	Black	136.76	17.13	21
			White	141.42	21.06	12
			Total	138.45	18.47	33
		Total	Black	138.41	18.50	32
			White	145.21	20.95	19
			Total	140.94	19.53	51
	Total	Present	Black	137.28	23.21	18
			White	149.50	20.20	20
			Total	143.71	22.25	38
		Absent	Black	136.20	18.35	30
			White	134.64	23.11	22
			Total	135.54	20.30	52
		Total	Black	136.60	20.70	48
			White	141.71	22.78	42
			Total	138.99	21.41	90
Blue	Present	Present	Black	142.45	19.45	11
			White	129.77	19.51	13
			Total	135.58	20.12	24
		Absent	Black	138.50	19.53	14
			White	126.13	14.67	15
			Total	132.10	18.01	29
		Total	Black	140.24	19.19	25
			White	127.82	16.86	28
			Total	133.68	18.89	53
	Absent	Present	Black	140.80	24.34	15
			White	138.78	16.49	18
			Total	139.70	20.12	33
		Absent	Black	143.78	24.35	18

			White	133.59	17.61	17
			Total	138.83	21.66	35
		Total	Black	142.42	24.01	33
			White	136.26	17.00	35
			Total	139.25	20.77	68
	Total	Present	Black	141.50	22.00	26
			White	135.00	18.08	31
			Total	137.96	20.04	57
		Absent	Black	141.45	21.18	32
			White	130.09	16.48	32
			Total	135.78	20.21	64
		Total	Black	141.48	21.91	58
			White	132.51	17.32	63
			Total	136.81	20.08	121
Total	Present	Present	Black	137.83	22.30	18
			White	139.04	21.86	26
			Total	138.55	21.79	44
		Absent	Black	137.09	21.12	23
			White	126.28	18.41	25
			Total	131.46	19.80	48
		Total	Black	137.41	20.84	41
			White	132.78	21.05	51
			Total	134.85	20.97	92
	Absent	Present	Black	141.12	22.69	26
			White	142.40	18.28	25
			Total	141.75	20.45	51
		Absent	Black	140.00	20.79	39
			White	136.83	19.15	29
			Total	138.65	20.02	68
		Total	Black	140.45	21.40	65
			White	139.41	18.79	54
			Total	139.97	20.18	119
	Total	Present	Black	139.77	22.33	44
			White	140.69	20.06	51
			Total	140.26	21.03	95
		Absent	Black	138.92	20.43	62
			White	131.94	19.38	54
			Total	135.67	21.16	116
		Total	Black	139.27	21.14	106
			White	136.19	20.10	105
			Total	137.74	20.64	211

APPENDIX I

Means and standard deviations for the interaction of
Font color x Background color for Part 2 of the CUSE

Font color	Background color	Mean	Standard deviation	N
Red	Black	136.60	20.07	48
	White	141.71	22.78	42
	Total	138.99	21.41	90
Blue	Black	141.48	21.91	58
	White	132.51	17.32	63
	Total	136.81	20.08	121
Total	Black	139.27	21.14	106
	White	136.19	20.10	105
	Total	137.74	20.64	211

APPENDIX J

Institutional Review Board Approval

Oklahoma State University Institutional Review Board

Date:	Tuesday, August 06, 2013
IRB Application No	ED13140
Proposal Title:	Variations in Online Survey Design
Reviewed and Processed as:	Exempt

Status Recommended by Reviewer(s): Approved Protocol Expires: 8/5/2016

Principal Investigator(s):

Celina Mendoza
1147 Sue lane
Choctaw, OK 73020

Dale Fuqua
444 Willard
Stillwater, OK 74078

The IRB application referenced above has been approved. It is the judgment of the reviewers that the rights and welfare of individuals who may be asked to participate in this study will be respected, and that the research will be conducted in a manner consistent with the IRB requirements as outlined in section 45 CFR 46.

■ The final versions of any printed recruitment, consent and assent documents bearing the IRB approval stamp are attached to this letter. These are the versions that must be used during the study.

As Principal Investigator, it is your responsibility to do the following:

1. Conduct this study exactly as it has been approved. Any modifications to the research protocol must be submitted with the appropriate signatures for IRB approval. Protocol modifications requiring approval may include changes to the title, PI, advisor, funding status or sponsor, subject population composition or size, recruitment, inclusion/exclusion criteria, research site, research procedures and consent/assent process or forms.
2. Submit a request for continuation if the study extends beyond the approval period of one calendar year. This continuation must receive IRB review and approval before the research can continue.
3. Report any adverse events to the IRB Chair promptly. Adverse events are those which are unanticipated and impact the subjects during the course of this research; and
4. Notify the IRB office in writing when your research project is complete.

Please note that approved protocols are subject to monitoring by the IRB and that the IRB office has the authority to inspect research records associated with this protocol at any time. If you have questions about the IRB procedures or need any assistance from the Board, please contact Dawnett Watkins 219 Cordell North (phone: 405-744-5700, dawnett.watkins@okstate.edu).

Sincerely,

Shelia N. Kennison

Shelia Kennison, Chair
Institutional Review Board

VITA

Celina Ann Mendoza

Candidate for the Degree of

Doctor of Philosophy

Thesis: VARIATIONS IN ONLINE SURVEY DESIGNS

Major Field: Research, Evaluation, Measurement, and Statistics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy Educational Psychology, option in Research, Evaluation, Measurement, and Statistics at Oklahoma State University, Stillwater, Oklahoma in May, 2014.

Completed the requirements for the Master of Arts in Psychology at the University of Central Oklahoma, Edmond, Oklahoma in 2007.

Completed the requirements for the Bachelor of Arts in Psychology at the University of Central Oklahoma, Edmond, Oklahoma in 2005.

Experience:
Research Assistant, Aging Services Division, Oklahoma Department of Human Services, November 2013-present
Graduate Research Assistant, Oklahoma State University, November 2011-August 2012

Professional Memberships:
Golden Key National Honor Society