A DATABASE AND JAVASCRIPT PROGRAM TEST OF INTERNET SURVEY METHODOLOGY: A CLOSER LOOK AT RESPONSE RATES AND DROP OUT RATES.

By

CHING-GUO WU

A thesis submitted in partial fulfillment of The requirements for the degree of

MASTER OF ART IN COMMUNICATION

WASHINGTON STATE UNIVERSITY
Edward R. Murrow School of Communication

MAY 2003
To the Faculty of Washington State University:

The members of the Committee appointed to examine the thesis of CHING-GUO WU find it satisfactory and recommend that it be accepted.

________________________________________
Chair

________________________________________

________________________________________
ACKNOWLEDGEMENTS

This thesis would not have been completed without the invaluable assistance of many people. Special acknowledgements are for Dr. Joey B. Reagan, the committee chair. He has always provided comments and suggestions crucial to this research. I am grateful for his ful assistance, patience, support, and encouragement. I would also like to acknowledge Dr. Erica W. Austin and Dr. Bruce E. Pinkleton, for their enthusiastic insightful, and constructive comments.

I would like to thank my dear friend Heather M. Crandall who helped me proofread and edit this thesis. I also want to thank many faculty and friends at WSU for their support and encouragement which has helped get me through this project. And friends in the Center for Teaching, Learning and Technology at WSU not only helped me develop the knowledge I needed to finish this research but also provided me with a lots of valuable ideas about Internet survey research.

Finally, I want to express my gratefulness to the precious people in my life: my parents and my wife Hao-Yu, who always makes me laugh and stands by me no matter what. Without them, I would never have made it as far as I have.
A DATABASE AND JAVASCRIPT PROGRAM TEST OF INTERNET SURVEY METHODOLOGY: A CLOSER LOOK AT RESPONSE RATES AND DROP OUT RATES.

Abstract

by Ching-Guo Wu, MA
Washington State University
May 2003

Chair: Joey Blaine Reagan

The Internet has been a popular survey platform in recent years. But the use of the Internet to deploy online surveys also has some problems. The most common problem reported by previous Internet research is the low response rate. It is a value we can easily obtain while conducting an online survey, but of interest to us here are the hidden meanings behind the number. The purpose of this research is to develop a method to better analyze the response rates to see if there is an approach which will improve the Internet survey.

To do this, the process of taking an Internet survey was first partitioned into four steps. Each step represents a situation that occurs while answering an online survey. A client side data collecting program was designed in this study to collect responses simultaneously as respondents answered the survey. Responses and corresponding time were collected once an answer was made on the survey page. With the new data collecting technique, all responses a respondent made were faithfully recorded along with
the time the respondent made the response. The extra information such as drop out point, response time spent on each question, etc … helped us better understand the Internet surveys.

By testing response rates and drop out rates on different survey invitation methods, reward systems and survey presentation styles in this study, the result showed that the drop out rate is important when we were able to draw more respondents. It also demonstrated that different survey invitation methods and rewarding systems have an effect on motivating respondents to start an Internet survey whereas survey presentation styles have only a slight influence of retaining respondents to the end of a survey.

The most important contribution of this study is that it validates the program we developed for observing respondents’ responding behavior. Some future research opportunities are suggested, and with the technique we developed, we now have a better chance of understanding how to make the new survey platform more useful.
TABLE OF CONTENTS

| ACKNOWLEDGMENTS                        | iii |
| ABSTRACT                               | iv  |
| LIST OF TABLES                         | vii |
| CHAPTER                                |     |
| 1. INTRODUCTION                        | 1   |
| 2. LITERATURE REVIEW                   | 2   |
| 3. RESEARCH QUESTIONS AND HYPOTHESIS   | 27  |
| 4. METHODOLOGY                         | 36  |
| 5. RESULTS AND ANALYSIS                | 53  |
| 6. DISCUSSION                          | 77  |
| BIBLIOGRAPHY                           | 89  |
| APPENDIX                               | 96  |
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>List of grouping categories of samples tested.</td>
<td>48</td>
</tr>
<tr>
<td>2.</td>
<td>Response results grouped by invitation</td>
<td>56</td>
</tr>
<tr>
<td>3.</td>
<td>Response results grouped by rewarding method</td>
<td>60</td>
</tr>
<tr>
<td>4.</td>
<td>Response results grouped by survey format</td>
<td>63</td>
</tr>
<tr>
<td>5.</td>
<td>Cumulative time spent on survey by survey presentation style.</td>
<td>66</td>
</tr>
<tr>
<td>6.</td>
<td>Time spent on each question item by survey presentation style.</td>
<td>68</td>
</tr>
<tr>
<td>7.</td>
<td>Time spent on each question item by survey completeness.</td>
<td>70</td>
</tr>
</tbody>
</table>
Dedication

This thesis is dedicated to my father, who educated me to be who I am.

Hope his soul rest in peace.
CHAPTER ONE
INTRODUCTION

As the newest medium, the Internet has become an increasingly popular form of data collection as well as a favorite topic of research for both academicians and industry researchers. The trend is that the Internet is becoming one of the most important resources for people seeking information, spreading messages, and even communicating with others online. How to better utilize this tool for conducting surveys and collecting response is also a valuable topic for social science research and is the major motivation for this study.

In order to replace the method that has been used for a decade with the Internet, some questions still need to be answered. For example, will the previous experience with traditional survey method be applicable to this medium of new technology? Are there factors that do not influence the traditional survey approach that emerge in Internet surveys? Our general interest was to investigate those uncovered factors that could influence the effectiveness of online survey methodology, and also to find out how these factors interact with respondents, and affect the results of collecting Internet responses.

In summary, the goal of this study was to exam what features or functions could be used with questions to enhance the intention to respond or to relieve the perception of responding as a burden! Furthermore, we wanted to know what style of survey design and what kind of invitation method or reward system would yield better response rates? By testing the hypotheses, we came to a more complete understanding of conducting a meaningful online survey.
CHAPTER TWO
LITERATURE REVIEW

Research on Internet surveys reveal that many researchers and practitioners are interested in learning more about the multitude of Web users in order to understand how this new technology can manage complex questionnaires more quickly, flexibly, and inexpensively than conventional survey tools. Some reasons is geared toward understanding at what level we can trust the data collected from the Internet (eg. Witte, Amoroso & Howard, 2000; White et al, 2000; Yoffie, 1998). No matter what the answers are, the interactivity and the ability to reach mass Internet users without delay already makes this new tool a terrific medium to publish instant information like breaking news, sports score, etc… (a possible replacement to broadcast media), and also to retrieve feedback from Internet users or audiences (an optional interactive channel to telephone, mail…). In addition, by developing an interactive web application, it will also provide online survey administrators with the ability to automate the whole survey process, including briefing, displaying information, interviewing participants, providing help, and recording responses, making the Internet an optimal medium. Because the Internet can be customized to model what human interviewers do during surveys using traditional methods, the survey procedure can be unified, hence all actions performed by the Internet survey application can be standardized. It not only guarantees that treatments for every participant are the same, but also eliminates human error that might be caused by the interviewers or by the interaction between interviewers and respondents. Characteristics used to describe other media like, “fast access, reliable, verifiable delivery, and ease of feedback” can all be used on the Internet (James, Wotring, & Forrest, 1995, p.47).
Because of this, researchers in the late 90’s predicted that the Web and the other new electronic technologies would soon become the “prime” survey vehicle (Kaye & Johnson, 1999).

Besides the characteristics and advantages posed, researchers are more interested in the quality of the data gathered online. Some studies have shown that the data collected online yield similar results to traditional paper-and-pencil methods (Best, Krueger, Hubbard & Smith, 2001; Yoffie, 1998) or need to be statistically adjusted (Witte, Amoroso, & Howard, 2000), and also suggest ways to improve response rates, as well as survey design and sampling quality (Couper, Traugott, & Lamias, 2000; Crawford et al, 2001). Additionally, researchers are still arguing about the low response percentages and non-controllable and non-representative respondents from the Web data (Gjestland, 1996; Gorman, 2000). What is still undiscovered is what makes these methods and the performance of these methods different. Without further validation, adopting the technology as a valid research method lacks a firm foundation. It is dangerous to use without knowing more about it. What can we do to improve the performance of online surveys in terms of representative respondents and quality data? Is there any technique that can help us measure the factors influencing a participant’s decision to respond to an online survey? What and how do these factors interact with Internet survey methods? How can technology help us improve the quality of online research? These are the questions this research attempts to examine.

Is the Internet the future?

Before trying to answer this question, let us take a look at some facts about the Internet. Despite media reporting the success of the Internet, there are more interesting
questions about the number of people actually browsing the Internet, what information are they seeking, and how are they sending and receiving electronic mail online? Of course, with the existence of the Internet, we can look for the answer by typing in the question in any online search engine. Thousands of answers will appear promptly from the virtual world on the other side of your fingertips. Based on a combination of sources, (Nielsen/NetRating, IntelliQuest, Nua Internet survey), the estimated number of people online worldwide reached 580.78 million in May 2002; about 9.57 percent of the total population worldwide. If we only focus on the number in the US and Canada, then it is 165.75 million; 59.1 percent of the total population up to April 2002. This 60 percent doesn’t provide enough information as to why the Internet research became such a popular field. But, if we know that the rate grew 15% per year in the last two years (the online population grew from 122.8 million to 164.4 million, 45.04% to 59.86%, in the year 2000, and from 79.4 million to 122.8 million, 29.3% to 45.04%, in the year 1999, and then slowed down due to the economic decline after the year 2000), we should not be surprised. It is a rapidly growing community. Of course, it’s not just the action at the user side that makes the Internet so popular. For example, when the telephone system was adopted, the value of communication technology increased as the number of people (or amount of information) joining the system also increased. As existing business and media move toward the World Wide Web, the online environment becomes more valuable to the users. Similarly, as more users connect online, more business opportunity will also appear for the information industry.

AT&T, the world’s largest long-distance carrier, made itself the largest cable television system in the United States by acquiring two large cable TV operators, TCI in 1998 and MediaOne in 1999 by outbidding Comcast, as part of its strategy to be an
integrated provider of TV, data, and phone service. In the year of 2000, Time Warner made a similar move by acquiring AOL. Finally, in 1995 Microsoft, the software giant of the world and also one of the major Internet service providers (ISP) in the United States, formed its strategic alliance with NBC, one of the nation’s largest broadcast companies, to extend their influence on both the broadcast and the Internet mediums. Microsoft also bought Hotmail, the largest web-based e-mail service provider, in 1998, and recently, it is trying to acquire Google, the most popular search engine in the world to become probably the most powerful enterprise in the Internet dynasty. Moreover, Comcast, who also received an investment from Microsoft after losing its bid on MediaOne to AT&T, finally merged with AT&T broadband in December, 2001. These moves and strategic alliances of these giant companies that try to control multiple media channels are not just coincidences. This phenomenon indicates that these enterprises are dedicated to controlling as many channels as possible.

On the other hand, the improvements on the technological side, such as wider bandwidth and faster hardware, makes the idea that mates the computer with television, radio, telephone, and all other communication and information devices highly possible. As long as the revolution in technology continues, the users will not stop coming. By reading the history of movements in the Internet industry, we can almost foresee that the effort to integrate various existing channels into one place will eventually be achieved. The Internet may not be able to replace current media, but we will definitely see it playing a formidable role in the future.

Using the Internet as a survey tool

With the burst of the online population, the Internet seems promising to a lot of
researchers and practitioners as the best choice for conducting surveys in the next generation over traditional mail and telephone methods. People working on survey methodology research are trying to find a cheaper, easier, and more effective place to administer their surveys and find that the Internet meets those criteria. Additionally, the growing speed of computer adoption also motivates faster technological development and application innovations. Implementing the innovations can also help generate more user-friendly and more function-rich Internet-based tools and add-ons for better interaction with users.

Today, conducting surveys on the Internet can not only be managed quickly, conveniently, and inexpensively, but also eliminates the need for interviewers or synchronous interaction (Mavis & Brocato, 1998; Schaefer & Dillman, 1998). The Internet also has the capability of experimentally manipulating stimuli, displaying audio and video, and holding live interaction or conversation between participants or between participants and facilitators (Hewson, Laurent, & Vogel, 1996; Smith & Leigh, 1997). For example, a survey investigating a person’s knowledge about sexually transmitted diseases might contain questions or pictures that would embarrass both respondents and interviewers. In such a case, the use of human interviewers might confound the data. But if question branching or interactivity is required, human factors would become significant in this survey. By adopting the Internet, the ability to provide automatic question branching, and also to provide some visual stimuli like displaying pictures or even streaming video eliminates the need for human interviewers in some types of surveys. Hence, the human error can be reduced. Furthermore, the use of the Internet allows the survey or experiment to be conducted from any location, from a private survey room or even in a respondent’s home, and at any time. This could help reduce a respondent’s
stress of answering personal or sensitive questions. It might also ease a respondent’s
defensiveness and enable the retrieval of more accurate information.

The list of examples like those mentioned above that show advantages of using
the Internet as a survey platform goes on and on. With so many advantages of the Internet
as a research platform, we will probably find more and more researchers using web
surveys (Buchanan & Smith, 1999; Davis, 1999; Witte, Amoroso, & Howard, 2000),
e-mail surveys (Couper, Blair, & Triplett, 1999; Schaefer & Dillman, 1998), Internet
facilitated experiments (Crawford, Couper & Lamias, 2001; Couper, Traugott & Lamias,
2001) and other Internet-based methods to perform surveys or experiments. The basic
assumption about administering surveys online is that, with the help of the Internet,
millions of Internet users could all be potential participants, including populations with
special characteristics (Binik, Mah, & Kiesler, 1999; Jackson, Ervin, Gardner, & Schmitt
N, 2001). Based on this assumption, the Internet has been used intensively in political
research like voting tendency (Taylor, H., Bremer, J., Overmeyer, Siegel, & Terhanian,
2001; Traugott, 2001), marketing research, psychological research (Buchanan & Smith,
1999; ), reliability and methodology research about the Internet itself (Bradley, 1999;
Couper, Blair, & Triplett, 1999; Couper, Traugott, & Lamias, 2001; Crawford, Couper, &
Lamias, 2001; Davis, 1999; Lang, 1996; Mavis, & Brocato, 1998), ethical issues of
conducting Internet surveys (Binik, Mah, & Kiesler, 1999), and also the qualitative
research that investigates the quality of focus group discussion in a web-based
environment (Schneider, Kerwin, Frechtling, & Vivari, 2002).

Currently, most surveys conducted via the Internet are just electronic versions of
questionnaires used in the paper-pencil based paradigm. They look exactly the same with
the exception of where they are displayed. Researchers treat this channel as another place
to present questionnaires, and take advantage of a bigger user base to draw a higher response rate inexpensively and expect these responses to be generalizable. However, the problem of low percentage response rates has arisen in many self-selected Internet surveys. The lack of representativeness raises scholarly concern over using the Internet as a survey tool. If we look carefully at the current online research, we will find that most of those studies are nothing but traditional surveys translated into a web page format. They simply use the Internet as a cheap channel for Internet users to respond by chance or upon invitation. Compared to the convenience, the ability of the Internet to control survey access is not taken into consideration in most existing web surveys. In other words, current Internet research is not taking proper advantage of the new technology.

Another consideration is that people use the Internet for e-mail differently than they use telephones or traditional mail, and people use multiple e-mail addresses for different occasions. People not only change their e-mail address when they move, graduate from school, or when they change jobs. People also change addresses when they switch their ISP provider, find a better deal for bigger e-mail storage, or just want to use a “cooler” name. In a traditional mail survey, one address could usually represent the contact outlet for one household. So when surveys are sent to a hundred different mailing addresses, we can pretty sure that we are inviting one hundred different households or persons for the survey. Conversely, in the Internet world, this assumption cannot be made. People have as many e-mail addresses as they want. They might use one e-mail address for business, one for home, and many different ones for different Internet communities in which they participate. In a case like this, ten different e-mail addresses sampled might actually be the same person. On the other hand, people who register an e-mail address might not use it regularly or long enough for Internet researchers to perform meaningful
sampling. This causes the e-mail invitation not to reach the sample in time for a valid survey response, or not reach anybody at all.

Because of the dynamic nature of e-mail addresses, the way we choose a random sample for traditional survey methods using address lists or a phone directory is actually not feasible. The Internet allows us to put more information online for more access and extends the possibility of reaching more people, but at the same time, audiences use the Internet because of faster information retrieval and easier search criteria. Because of this audience perspective we know users will have a lower tolerance for the burden of participating in lengthy surveys (Crawford, Couper, & Lamias, 2001; Dillman, 2000), or in surveys that do not offer appeals and can stand out against conventional questionnaires.

Unlike the long history of paper-pencil based surveys, the Internet survey is just entering the learning stage for serious practical implementation in academic hypothesis testing. Diminishing the value of the Internet in survey research because of some sampling frustrations of online data collection is unnecessary. The assertion made by Crawford et al (2001) that traditional survey methods are losing the credibility by lower and lower response rates encourages us to further explore new ways and opportunities to use the Internet as another survey platform. We can say that when new technology changes the way people seek information, the way they respond to what they perceive will also change. Thus, public opinion researchers should move along with the communication revolution by designing and conducting surveys in a new way. History has shown this kind of movement repeatedly as radio, telephone and television gained popularity in the past. To investigate how we can improve the Internet as a survey platform, further understanding of both its advantages and disadvantages is necessary. In
the following section, we discuss some concerns about using the Internet as a survey methodology. We also will look at the advantages the Internet poses and consider if these advantages can be used to overcome the concerns.

Concerns, Advantages and Limitations

Let’s start with the concerns that have been raised regarding the use of Internet surveys. These concerns will be scrutinized in three categories, Internet coverage, sampling problems, and the participants’ psychological processes. Dillman (1978, 2000) indicates that there are three types of errors, non-coverage error, non-response error and sampling error, that usually affect the quality of Internet research. In the next three sections, we will cover those error types by the three categories.

1. Concerns

   Internet Coverage

   One very important factor that makes the Internet an attractive medium for public opinion and media researchers was the size of the online population. In current web-based technology related research, we find that most of the articles mention the exponential growth of Internet access and usage as one of their research motivations. But even though almost 60 percent of the people in the US and Canada have Internet access already, Internet user demographic research (National Telecommunications & Information Administration, 2000) and others (Yoffie, 1998) expressed concerns over the fact that Internet access is not universal. The non-coverage error might cause certain attributes from non-users to be systematically ignored by the Internet surveys. For example, the background of the Internet users might be different from the non-users, which might generate misleading conclusions about the general population (Best,
Krueger, Hubbard, & Smith, 2001). These studies also mention that Internet users have attained higher education, have a higher socioeconomic status compared to non-users, and are more likely to be young, white, married and wealthier compared with the population at large (National Telecommunications & Information Administration, 2000).

Other research found that the gender gap in Internet use for adults in the U.S. has disappeared, and trends toward equal access by education, income, and race have also been noted (Clemente, 1998; Glasner, 1999; Katz, 1997, Witte et al, 2000). It’s interesting that opposing findings about online coverage are generated within almost the same period of time by different studies. Some researchers might argue that results of early Internet studies were based on the obvious demographic differences between respondents online and those collected via the traditional platform. As more is learned about average Internet access, we will not be able to use mixed results as an excuse forever. One example is the famous Literary Digest Poll error at the early telephone adoption period. One past failure doesn’t mean that it will happen again in the future. If technology doesn’t stop going forward, neither should public opinion researchers. This research will try to investigate the problem from some other angles besides demographics.

Best et al (2001) suggests that the reasons Internet research can generalize to the larger population are based on two assumptions. First, the decision-making processes of Internet users are no different from those in the population at large. Second, representative samples of Internet users can be drawn. Both assumptions are important topics in Internet survey methodology research. They provide an outside view of efforts we can make to improve the quality of Internet research. The demographics of the Internet user might change as the Internet use becomes more popular. The problem is that we have no way to control this when planning research on the Internet. However, once
we have more knowledge about a respondent’s decision making process while participating or responding to online surveys and about better ways to draw samples from the Internet user population, we should be able to conduct better research by using more appropriate methods. Even with the current limitations of uneven Internet coverage, some specific populations are actually well covered by the Internet, such as PR and marketing professionals, computer enthusiasts, and software and hardware engineers. Coverage should not be the main concern if an Internet survey is targeting any of these populations.

If the coverage problem is not an obvious factor in Internet surveys over time, then other factors that might affect the success of Internet surveys and should be analyzed more closely. In the next two sections, we will focus on the other two concerns specific to Internet surveys.

**Sampling and data analysis**

Probability sampling is a crucial consideration in deciding on an appropriate platform for surveys. One reason is because probability sampling allows us to calculate the confidence with which we can generalize the results from the samples to a larger population using a statistical method. According to Babbie (1990), confidence in the representativeness of survey marginals require that every unit in the target population possess some chance of being selected so that the statistical likelihood of drawing each population unit can be computed. A survey is meaningful only if the sample selected can represent the whole targeted population in the survey.

Of course, we would not have any trouble generalizing survey results if we could collect data directly from the whole population. But in reality, what we can do is try to be sure that the sample we draw from is a fraction of the targeted population and can be
projected onto the whole population, which relies on all units in the target population having an equal chance of being selected. A series of steps have been developed to improve the process of sampling. First, we have to understand the survey population in order to choose a feasible sample frame and then select samples from this sample frame. These procedures are scientifically proven and have been implemented in survey research for decades. The most popular and the most trust-worthy method is probability sampling. The advantage of probability sampling is that when samples of a given size are drawn repeatedly, corresponding samples estimates will form distributions from which the true population parameters can be derived.

To distinguish between probability and non-probability sampling, probability sampling is characterized by the fact that the sample is selected by chance, and the population members have a known, and sometimes equal, probability of being selected. The probability techniques most familiar to us include simple random, systematic, stratified and cluster. In contrast, non-probability sampling uses human intervention, and includes quota sampling, judgment sampling and convenience sampling.

Since probability sampling has been the major sampling method in social science research, Internet researchers also have adapted these traditional sampling techniques in the online environment. The question of whether a representative sample can be drawn from the virtual population became important for operating Internet surveys. Studies found that only diverse, non-representative samples of Internet users could be generated (Best, Krueger, Hubbard, & Smith, 2001). They analyzed three different sampling units that are frequently used to construct sampling frames of the Internet population: listed e-mail addresses, electronic subscription groups, and Web sites with heavy traffic, and found no support for the assumption that a representative sample of Internet users can be
drawn no matter which selection criteria was used. The problem, they stated is not in the way samples are being drawn. They indicated that the sampling frame from all three categories cannot be designed to ensure that each Internet user possesses an equal chance. Sampling frames that use list e-mail addresses exclude numerous Internet users who do not appear in online directories. The Internet users who do not join newsgroups or mailing lists have no opportunity of being selected by the Internet survey conductors. And the amount of traffic on a given Web sites changes from time to time. A Web site with more traffic doesn’t mean that more Internet users will notice the existence of a survey link posted on that site, or that those who notice the survey will click on the link and opt to visit a survey page from that Web site.

Yoffie (1998) also argued that online research is not “representative” because the “self-selection” nature of the online survey overly depends on the chance of visiting some specific sites. Furthermore, it is difficult to tell the real number of distinct Internet survey participants who complete the Internet survey. Thus, the data quality is compromised. The question becomes whether three responses really represent three different people or is one specific person actually responding to the survey multiple times? Although most studies provide weak evidence in relation to proving that the samples drawn from the Internet by self-selected responses can be treated like those drawn using RDD or other conventional probability methods in a non-Internet environment, they also pointed out that parallel studies show similar results between online and traditional research methods (Yoffie, 1998; Witte, Amoroso, & Howard, 2000).

The contradictory findings detailed above showing similar results derived from research using traditional probability sampling methods and online non-probability self-selected sampling methods is worth some in-depth study. If we define the
respondents who self-selected themselves into the Internet survey as “non-representative” samples (due to the nature of the online sampling method) the finding that “non-representative” respondents and “representative” respondents yield similar results in data analysis becomes interesting. It is not appropriate to make the judgment that the representative sample or a proven random sampling mechanism is not necessary in the Internet survey on this comparison alone. Instead, there might be other factors that cause similar results using different sampling methods. For example, is the low response rate in traditional survey methods, especially in mail surveys, taken into consideration? Is it a sound decision to treat on-line respondents as non-representative because of the “self-selection” nature and ignore the fact that the assumed representativeness from traditional random sampling methods poses the same possible analysis error because the representativeness is also harmed by the same low response rate? Or should we say that some factors are ignored during the sampling and testing process, and those factors are the reason for the contradictory finding?

Smith (1997) pointed to this problem of comparing data from web surveys and random sampling surveys with low response rates. According to Smith (1997), “This issue is not unique to web-base survey research.” What about the assumption that individuals who choose to complete an Internet survey (who self-selected themselves to respond to the Internet surveys) are actually among the same people willing to complete a non-web-based survey after being randomly selected by traditional sampling methods? Maybe voluntary participants online are more interested, informed, and concerned about the survey topic, and they typically hold viewpoints that are stronger and more extreme than those held by other individuals (Asher, 1995; Rosenthal & Rosenow, 1969; Wu & Weaver, 1997), similar to those respondents who decide to complete surveys in traditional
To examine those assumptions, some more considerations about respondents’ psychological processes are necessary. First, are there undiscovered factors in the decision-making process that occur while a person is responding to an Internet survey? Second, is there a way to analyze those hidden factors in the respondent’s control after the surveys is released? These questions will be discussed in the next section on the respondents’ psychological process.

**Psychological process**

Once an Internet survey is published on a website or sent to the selected participants, the survey conductors can only wait until they come back. Actually, if we scrutinize the whole process, no matter how much time has been spent on designing and distributing the questionnaire, the only thing that distinguishes success from failure of the prepared survey is the question of whether it will bring back what we want to know “accurately”? We can control the content of the questionnaire at the design step, and we can control the data analysis process after we gather responses, but we can hardly say that we can control a participant’s (a sample) reaction to completing it during the testing step.

Most studies (both mail and Internet survey’s) suffer from the problem of low response rates according to previous studies. Response rates in Internet surveys have been reported to be as high as 70% (Brennan & Hoek, 1992) and as low as 0% (Pradhan, 1999) with around 30% as an average for a better execution. Research also finds that comparing the response rates between traditional mail surveys and Web surveys varies. Some found higher response rates in e-mail surveys (Brennan & Hoek, 1992), some found nearly equal response rates between the two methods (Bachmann, Elfrink, & Vazzana, 1996;
Mehta & Sivadas, 1995), whereas others have reported that the web surveys returned less (eg., Kittleson, 1995; Tse et al., 1995). No matter how we try to eliminate the differences in coverage and sampling, low response rates are still a threat to Internet research. With less than a 30% response rate, it still might not matter that we can choose perfect samples. This is because almost twice of the non-respondents may still be systematically lost for various reasons during the survey, and 30% of the valid response can easily be rewritten if only half of those non-respondents have responded in another direction. Thus the non-response error is as serious as the non-coverage error and the sampling error.

By reviewing previous research conducted online, we find that even though most Internet surveys suffer from low response rates, the non-response error is actually getting less attention than the other two types of errors in the survey planning stage. It seems that little effort has been made to reduce the non-response rates prior to starting online survey research. Most studies that used the Internet as a survey platform and obtained low response rates only stated this fact and explained that low response rates are commonly found in most Internet surveys nowadays. Some did make the effort to adopt the suggestions from the Internet survey methodology studies and in some cases achieved improved response rates, but the full mechanism behind it is still uncertain. Some tried to justify their findings by comparing results gathered from the Internet with those that used a traditional paper-pencil method. But as was argued in the previous section, this kind of ignorance on low response rates is very dangerous if social scientists continue to use the Internet as a formal survey tool before thoroughly understand the factors that might cause errors in survey methodology.

With the Internet access rate approaching 60% in the US, we can say that the real challenge to Internet research is figuring out how to improve the response rate of Internet
surveys. Crawford, Couper, & Lamias (2001) named one reason that web-based e-mail surveys appear to be attaining lower response rates compared to mail surveys may be because there is currently little information on effective strategies for increasing the response to Internet-based surveys. Some suggestions have been made to improve response rates for web surveys, like mentioning the survey length in the e-mail invitation (Crawford et al., 2001), sending reminders after the invitation (Dillman, 1978, 2000), and using graphic progress indicators (Couper, Traugott & Lamias, 2000). As you can see, most suggestions are based on previous knowledge with traditional paper based surveys, and only one of these uses the capability the Internet provides. Some studies have tried to test the effectiveness of these suggested techniques; however, none of them are repeatedly achieving improved response rates for Web surveys. In general, researchers suggest that the difficulty of the response process is the major reason for low response rates or “the more burdensome the task, the lower the response rate” (Babbie, 1990), but Crawford et al (2001) argues that there are a large number of other factors not yet being studied. These undiscovered factors potentially involved in the respondents’ decision making process to respond or not is the focus of this study. In the following sections, we will try to uncover the possible factors by introducing “decision making analysis” and the “dropout rate”.

### Decision making analysis

If we analyze the whole decision making process involved in responding to online surveys, we can separate it into four major steps (see figure 1). Each step represents a breakthrough in the survey proceeding. First, the respondent needs to perceive the
information about the survey, and then he or she needs to make the decision to take this survey for the whole procedure to continue. Once this first decision is made, the respondent begins to answer the questions until finished, and then submit the survey. If we follow the logic of the four-step survey taking process, we can see that respondents will need to make three decisions to complete this survey. At each step, any factor may at any point influence the respondent to give up. He or she might never start the survey (step 1), might stop after opening the survey and before answering it (step 2), might stop at some point during answering (step 3) due to time constraints or being reluctant to answer sensitive questions, and he or she might not submit it after finishing it (step 4). Since there are so many possibilities for quitting the process, how can we determine why a respondent failed to complete the Internet survey by referencing only the response rates?

The delivery method is one factor we can control for entering step 1. We can get some idea about factors that influence continuing to proceed by analyzing the decision flow. Basically before they actually enter the first step, we have to make sure that the survey information actually reaches the respondent. But just because the information is delivered doesn’t mean that it won’t be ignored, thus the question in the Internet survey methodology here becomes what can we do to have the information we delivered not be ignored? This is the factor we want to control while administering the survey.
Figure 1. Psychological process on continuing response during testing.

Once respondents perceive the survey information (enter the first step), we need to increase the percentage of respondents who actually start answering the survey. If we can have the respondents notice the information we delivered, how can we attract them to being responding? At this moment, all the information the respondents have about the survey is the invitation, hence, how to design the invitation letter (e-mail) and what to put into the first impression material they receive is what we can do to control the
psychological process flow at this stage. Usually some *reward* will be promised for respondents who answer the survey, and Musch and Reips (2000) show that the survey completion rate increases if some reward is promised. It might one the place we can put in the effort.

When a respondent starts taking the survey, there are still chances that he or she will stop at some point without reaching the submit button. In this step, a lot of factors can influence their intention to complete the survey. Basically, every question poses some influence, and different survey styles might also give the respondents different feelings about how difficult answering the survey is. Since survey questions are usually stiffer than survey style, the *survey presentation style* is something we can work on in this stage.

**Adopting “dropout rate”**

Comparing samples drawn simultaneously using the Internet and probabilistic telephone methods, research shows that the psychological mechanisms underlying common political decisions do not differ between Internet users and the population (Best et al 2001). This finding gives us some confidence in relation to the data entry by the same respondent on different survey platforms. Once they take the survey, they will put the same answer no matter what platform survey is presented. As we stated above, the response rate is not a good enough index for us to discover factors that influence a respondents’ survey taking process. Some researchers start to measure the *dropout rate* along with the response rate, to help them further justify their analysis. Studies indicate that participants in an unmotivating, boring, or very difficult experimental condition are very likely to drop out of the Internet experiment. Musch and Reips (2000) also state that
the 34% average dropout rate of Internet experiments actually range from 1% to 87%.

Frick, Bächtiger, and Reips (2001) suggest that factors such as the order of personal information and financial incentive questions will influence the dropout decision after the participants self-selected into the study. If we say the response rate indicates the percentage of respondents that go from step 1 all the way through step four, then the dropout rate can give us some indication of the percentage of respondents that go from step 2 to step 4. We can show this by using the following formula:

\[
\text{Total participants} = \text{number of non-survey-takers} + \text{number of drop out} + \frac{\text{number of response}}{(A) + (B) + (C)}
\]

The three groups of participants were differentiated by the steps they did or did not reach. Respondents in group A never reached the second step, respondents in group B dropped out at either step 2 or step 3, and respondents in group C completed the survey. Since we know the answering (testing) occurs during step 3, we will know if we can reduce the number in group A and group B, and can improve the research quality as well as generate more reliable results. It is reasonable, as Reips (2002) suggests, and good practice to always report the dropout rates of Internet-based research in order to better analyze the data and to try to understand what confounds motivation during research participation.

2. **Advantages of Internet surveys**

   Besides some concerns we discussed above that bother researchers who try to
conduct surveys on the technology based platform, there are actually many advantages to this platform. Some of these advantages are not readily apparent. To take full advantage of the Internet as a tool for conducting surveys, researchers also need to be prepared for the new technology in the Internet paradigm.

Cheap, fast, and easy to administer

New technologies are always integrated into public opinion research because of their ability to reach more people easier, faster, and at a lower cost. The Internet makes it possible to collect thousands of interviews in a day without human interviewers, which saves both time and money. Some researchers and practitioners think “they are a free, fast and entertaining way to get the audience involved (Palser, 2000).” Compared to traditional survey platforms, you only have to post one copy of the survey one time on the World Wide Web (WWW), and it all happens in a couple of “mouse clicks”. No phone calls, no postage, and no delivery delay. Amazed by the large amount of responses collected in a short period of time compared with traditional human mediated surveys, early researchers who adopt the Internet as their major survey platform are interested in the low cost and large numbers of responses. In recent years, as analysis and understanding of the Internet as a research tool emerged, data quality of online responses became a major issue in Internet research. Additionally, critics also try to clear up the myth of whether “online research is cheap,” and assert that poor and inaccurate online research may be cheap, but good online research will or could cost real money (Taylor, Bremer, Overmeyer, Siegel, & Terhanian, 2001). The bottom line is, that it can save survey researchers a lot of trouble if this method can also produce quality data.
Interactivity

Unlike television, radio, and mail that can only present information to audiences, the Internet is the medium that can distribute information, and also receive audience responses. Unlike the telephone, that only provides audio stimuli; the Internet can provide both audio and video. The most important feature of the Internet is that it provides real time interactivity between the web sites and the Internet users, or even between Internet users and Internet users. These characteristics allow us to model whatever our interests in the real world, like conversation, real time help, or a video conference. Surprisingly, none of these features have been used in current Internet surveys yet, but surely, we will see their implementation in the near future.

Flexible and programmable

Indeed, the ability to program an Internet survey to perform necessary stimuli, randomize question order, check response, automatically branch during progress, display audio and video while needed, and provide different interactive effects or help messages, makes the Internet a possible “dream” research platform compared to the traditional paper-and-pencil, human intervening paradigm, especially for administering an experiment. One significant difference between the computer-based interface and the human interviewer is that the computer-based platform always does what it is meant (programmed) to do, without exception, unless the server or system is down (see limitations below). Emotion, bad mood, approaching schedule, etc will no longer affect the data collection. This is the reason the computer was first introduced in CATI (Computer assisted telephone interview) and CAPI (Computer assisted personal interview) forms in survey research as it assists interviewers and facilitates the interview process. In
this way, mistakes made by interviewers can be reduced, or eliminated. In addition, although the Internet has the ability to be customized as a tool that meets all kinds of requirements and performs optimal interactivities, it does require some level of programming skill to be able to administer both the server side control and the client side scripting effect. In addition, to the appeal of mass exposure and online survey access, some studies have begun to explore how this technology can enhance data quality and reduce duplicated responses by adopting the server side authentication technique to request and process access codes (Crawford, Couper, & Lamias 2001) and pin on survey access control (Heerwegh & Loosveldt, 2002). On the other hand, Couper et al (2000) tried to understand how the technology could be utilized to improve response quality by providing visual effects that indicate to respondents’ their progress in the survey as they proceed toward the end.

3. Limitations

Despite that the technology allows researchers to simulate the methods they are familiar with from traditional survey research, there are some limitations. Three access barriers (or compatibility issues) restrict the use of the Internet, and will reduce or even eliminate research design efforts that try to include these programmed effects. Of course, the physical access of the Internet limits reaching the real world general population.

Another barrier is the network access barrier, which might be limited by the firewall, the bandwidth, the government, the institution, or a system’s filter setting. For example, a well designed survey that includes a nice audio or video effect could become a disaster with a slow dialup network, or may even be treated as suspicious pornography and blocked by the parental programs. Software access is the third barrier. Executing most
Internet research usually involves an e-mail invitation which includes a link to the actual survey page or a web form for direct response. But respondents who use text mode e-mail clients like “pine” in the UNIX system will not be able to take advantage of this convenient function. Some effects created using client side script technology might only work on certain operating systems or application platforms. (Currently, the Microsoft invented VB script is still not supported by the Netscape browser years after being introduced.) Because of the software limitations, a lot of system technicians are facing the dilemma of whether they should adopt the latest technology and provide a function-rich system, or slow their pace and take care of the legacy. This situation will, of course, give Internet researchers more challenges when conducting their studies. The Internet might be a great tool, but it won’t fully achieve its potential unless researchers have enough knowledge about how to handle it and how to overcome its limitations. After all, it will save time collecting data, but will cost more in the research design stage in order to ensure meaningful data collection.

After reviewing both the advantages and the disadvantages of Internet survey methods, and understanding the current challenges of conducting a successful Internet survey, we can now look at the unknown factors underneath these fancy technologies. A couple of questions will be addressed in this study in order to develop the required knowledge for using the Internet as the survey tool of the next generation. Those research questions are listed in the next chapter.
 CHAPTER THREE
RESEARCH QUESTIONS AND HYPOTHESIS

Response Rate

One thing that keeps bothering Internet researchers are low response rates. Thus, it is important for Internet researchers to know at what degree results generated from Internet research can be used to support research hypotheses. First, the most obvious suggestion from previous studies is to rescue lower and lower response rates. Hence, the first research question we tested was:

**RQ1:** Is there a way to improve the response rates of Internet surveys?

Even though much research has been done using the Internet and treating randomly selected e-mail addresses as probability sampling methods, and asserting that there is no significant difference between the data collected via the Internet and the data collected using conventional probability methods. Few have tried to exam the real differences between these methods in terms of how the decision making process works while the respondents participate in an Internet survey. Frankly, without improving survey response rates in Internet research, the use of the Internet as a survey platform is still suspect. If the response rate can be improved, the next question would be:

**RQ2:** What can we do to better understand a respondent’s decision to not complete an online survey in hopes of improving the response rates? What method we can use to understand the factors that influence response rates?

We can approach answers to these questions by investigating respondents’
responses step by step according to their decision making process while answering surveys online. As was analyzed in previous sections, advantages of the Internet such as interactivity, programmability, and the ability to execute real time data collecting make this attempt easier than using traditional survey platforms.

Generally speaking, the conventional research methods, either web-based or in the traditional format, use the data produced from a procedure without much questioning of the decision making processes and the many possible undiscovered factors involved in the decision making flow. We cannot simply trust what we get without considering further what variables are ignored. To pretend that those variables have little influence during a respondents’ decision making process while interacting with surveys is also dangerous. One thing we can do to clear those doubts is to analyze the decision making process by cutting it into smaller pieces. We should then try to collect as many of the respondent’s movements as possible within each piece for further analysis. We can do this by using a traditional observation method.

We will likely encounter even more difficulties due to a combination of the Internet as a survey platform and the complexity of using human as the observers. More errors will be introduced when using human observers on observing human-Internet interaction. First, it is hard to be objective if we use people as observers. Second, with the use of people, human mistakes are difficult to measure during the observation. And finally, by observing with human observers other than the Internet itself, the appropriate probing points are hard to determine. Hence, missing the right point is highly possible if we concentrate on a fixed period, and wasting resources and time is unavoidable if we try to fully cover the observation. It is for these reasons, that the Internet is the best candidate for us to observe and analyze a respondent’s decision making while they are responding.
It brings us to the next question for this study:

**RQ3:** Do we have a computer-based method that can help us better observe respondents’ behavior objectively, so more factors that influence Internet surveys can be discovered?

**Samples**

The use of student samples has been criticized as not actually generalizable to the larger population (Best et al 2001, Lang, 1996; Sears, 1986). The fact remains that currently all students in the target university (Washington State University) have Internet access and a distinct e-mail address that can be retrieved from the school’s online directory, and is still a better population for a methodological study. We have the opportunity to test a respondent’s decision making process of answering an online survey on an averagely computer literate, homogeneous, Internet user population such as college students. Students are suitable as a target population in this research for a couple reasons: First, we have the real population (student and faculty e-mail directory representing each member of the school community) to draw a sample from, so the sampling error can be controlled. Second, because the university provides every student a chance to use the Internet (assigns e-mail addresses, provide Internet access in the library and the computer labs, and offer online learning environments), the coverage error can be avoided.

**RQ4:** How can a student population help us to develop an experiment testing the potential factors involving Internet survey research such as survey information delivery and reward methods?
Dropouts versus Non-responses

In this research, another important task is to know if the non-response rates (people who didn’t respond) in Internet research is different from the incomplete survey rates (including drop out rates) in order to understand if further analysis in a respondent’s drop out decision is necessary. In most previous survey research, only the response rate is available for data quality validation. Drop out behavior is unavailable in mail surveys or Internet surveys and no method has been suggested to collect it for interest alone. However, the non-response rates collected in previous research doesn’t guarantee that all those non-respondents didn’t try to respond at all. If the non-response rate is collected without considering the existence of dropouts contained in a large percentage of them, then the attempts to improve conducting Internet surveys have been misguided!

Since the two major types of error can be controlled, the non-response error can be investigated easily by collecting data at the point the respondent drops out while taking the survey and compare this to the non-response rates and drop out rates using the server side programming technique. Therefore, we might be able to analyze a respondent’s decision making process during the Internet experiment more accurately. If the drop out rate is as high as the non-response rate (a significant amount of people drop out of the survey in the middle of responding compared to the percentage of the sample that never start it), or even higher than it, than more effort is needed to reduce this dropout; otherwise, attracting people to begin the survey becomes more important.

RQ5: Does the dropout rate play an important role in the Internet survey? Or it is relatively non significant comparing to the non-response rate?

Survey Appearance and Form of Invitation
Besides the response rate and the dropout rate, we can analyze the respondents’ survey answering behavior to capture some potential factors that might play roles in influencing the survey response rate. We want to understand if the style of the Internet surveys (the way they appear, and the form they are presented on the conventional browser) will influence a respondent’s response or their decision to drop out. To know whether these factors can help us in implementing a better online survey, the following questions are also important in this study:

RQ6: Does the format that an Internet survey is presented in influence a respondent’s decision making process? If the same online survey is presented in different styles, would the survey styles influence a respondent’s intention to stop or finish the survey?

Furthermore, in conventional Internet survey practices, several methods that invite audiences or target population into an Internet survey like, e-mails, popup windows and advertisements are used, but how effective are these methods? These methods have different visual effects for respondents who receive them, and could also influence the psychological perception and decisions while responding to a survey. Thus, the answer to the next research question might help us develop a better invitation technique to draw more responses from Internet surveys.

RQ7: Will the invitation format influence participants’ willingness to join or complete an Internet survey? What about the content that we include in the invitation message? Can rewards help draw responses?

Based on these questions that are of interest to us, and the results from previous
research, we propose the following hypothesis for this study.

**Hypotheses**

The primary purpose of this research is to understand if there is any difference between the non-response rates and the drop out rates. If the drop out rates just determines a small fraction of all of the non-response rates, then we can say that the drop out rates do not play an important role in a respondent’s decision making flow while taking an Internet survey. From the analysis in the previous section, we know there are many steps involved in making a decision about whether to proceed or not after a respondent decides to take the survey and before they submit it. It is highly possible that any factor within this period could interrupt the response action. If this assumption is correct, the non-response rate reported in many Internet surveys might have included the dropouts and should be dealt with separately from the total number of non-responses. The first hypothesis we want to test in this study is:

**H1: The non-response rate will return different results in data analysis.**

It underscores the important role drop out rate plays in Internet surveys. If the hypothesis is supported, more effort should be put into finding the factors that discourage a respondent from completing an Internet survey.

Besides the questions in the questionnaire, survey design is another place that we can work on in order to get better responses in terms of response rate and data quality. From existing research, not much work has been done in the field of investigating the effect of survey styles, delivery methods or reward types that may influence a respondent’s intention of answering or completing an online survey, and further cause the
dropout or response rate to vary. Inherited from traditional methodology research, most of the Internet survey methodology studies still focus on comparing demographic differences, (Yoffie, 1998; Clemente, 1998; Glasner, 1999; Katz, 1997, White et al, 2000), sampling implementation (Best, Krueger, Hubbard, & Smith, 2001; Smith, 1997) and response rates instead of questioning the structure of the questionnaire they adopt from the traditional survey methods. Because of the characteristics unique to the Internet like the massive information available and less time and environmental stress during self responding, many new factors could interfere with a respondent’s responding by diverting their attention from the task they are working on.

Our analysis also indicates that under a respondent’s decision making flow, there could be many factors have not yet been discovered. Even though only a few studies can be referenced, many methods and techniques are developed and practiced in commercial or real world implementation. These techniques include using different invitation methods to attract audience attention, using different presentation styles, paging, and progress indicators or rewards to enhance their intention to complete an online survey. They are not only factors that are unique to Internet surveys; they could also be applied to the traditional methods. The reason that they did not get discussed before may be due to the fact that they are not significant factors to the traditional survey methods. But are they significant in the Internet era? The analysis above makes us believe that these factors might play more important roles in obtaining higher response rates from Internet surveys because of the Internet characteristics. The second hypothesis will test these factors in Internet surveys. Because different factors might influence responding behavior in different steps (some might have effect on more than one of the decision processes, and some might have only limited influence to one of the stages described above), the
hypotheses need to be tested separately into three sub categories as below.

**H2a:** The style of the Internet survey will influence a respondent’s decision-making procedure in completing an Internet survey (dropout).

**H2b:** The methods used to deploy the Internet survey will influence a respondent’s decision-making procedure in responding to an Internet survey (response).

**H2c:** The methods used to deploy the Internet survey will influence a respondent’s decision-making procedure in completing an Internet survey (dropout).

First, the respondents did not see the survey style prior to entering (responding to) the survey, so survey styles shouldn’t be a factor in their responding decision. We tested the survey styles primarily on its influence to a dropout decision. On the other hand, survey deploying methods might not only impact a respondents’ decision to respond, but also impact willingness to complete the survey. Hence, to test the influence on both steps in then responding process is necessary.

If any of these hypotheses is supported, then we know exactly where in the Internet survey responding process to inspect and research in order to get better representation from Internet surveys. If we know the weaknesses of the Internet survey conducting methods, it will be easier to figure out how to improve the survey platform.

Another important issue we tested in this study was to learn whether there is a technique that can help us examine the potential factors that might influence Internet survey completion. To investigate this, an embedded task-specific program was created to collect response interactively. Conventionally, the HTTP (Hypertext Transfer Protocol) protocol is the most popular technique in current online data collection. But responses
cannot be collected until respondents explicitly click the submit button implementing the basic function the HTTP protocol provides. This means that the respondent input might be lost if the respondent decides to stop responding to the survey before completing it. Thus, all incomplete response will look like non-responses during data analysis.

If we want to collect the drop out point information, or the time respondents spend answering each question, the traditional method most researchers currently use is no longer feasible. Some techniques are introduced to solve this problem by capturing the data immediately after every response using static connection created by either an installed application or a java applet. These programs open a connection socket at the server side in order to provide the specific service of collecting data simultaneously as the actions happen at the client side. But these techniques require huge overhead (installation of an application or a long time to download an applet into a local machine), which will cause the respondents to give up before even seeing the survey. The tool we designed is a small piece of JavaScript code we embedded inside the survey page. It can help us collect responses simultaneously without suffering the expensive overhead of big program downloading, and will also be tested in this experiment.
CHAPTER FOUR

METHODOLOGY

In this study we conducted an experiment observing a respondent’s responding behavior in the form of a regular Internet survey. To simulate the situation as respondents are taking an Internet survey, one set of questionnaires will be used. Since we are testing at what point in time a respondent drops out of the survey or the time a respondent takes to complete different survey formats, a respondent’s answer to questions are not the most important data we are collecting in this experiment. However, in order to make the respondents think they are actually taking an Internet survey, the content of the questionnaire and the method used to conduct it were dealt with carefully. It could be more helpful if the Internet questionnaire contained some questions that track Internet usage or personal information regarding the respondents, such as frequency of use, computer literacy and demographic data. The content of a paper-based questionnaire that has been used in previous Internet usage research conducted by Busselle, Reagan, Pinkleton and Jackson (1999) has been adopted based on the above reasons.

To investigate the respondents’ drop out and response rates while taking an Internet survey, an Internet experiment was designed. The main challenge in this experiment was utilizing the Internet's capability of observing a respondent’s decision-making behavior during the survey process. We need to be able to collect all the data and record the time it takes a respondent to make a selection or response to the survey without obvious delays or the need to refresh the survey page. In other words, we need a program that is able to start recording a respondent’s response (counted as a responded respondent) once the respondent enters the survey (corresponds to step 2 in
Like what was stated, not only should the response be collected after the respondent completes the survey and clicks the submit button, these responses should not be lost once they answer them. Besides the responses, respondents’ behavior during responding can also provide important information for survey analysis. Capability of tracking their responding process throughout the survey session is another key issue here. A technique is developed in this study for these reasons.

**Data collecting technology**

After three years of web application programming experience on an online learning system, there are visible limitations of HTTP protocol, which is used by most online surveys to collect information. This protocol which records whenever a respondent makes a selection faces some challenges. Like what was described in the previous section, the regular use of a web form allows us to collect data while respondents submit the form, which causes the page to be refreshed, repositioned, and is significantly delayed if the dial-up connection is used on the client side. All of these side effects are not what we want to see when trying to keep the respondents answering the survey. So, even though some ideas or techniques have been used to create static communication connections between a server and a client, the requirement and overhead will not help us in this research.

To solve the conflict between introducing a new technology and the ease of its use, a new technique was developed to use a frame set and transmit only a small amount of data at each submission. This is based on the HTTP architecture. There might still be some slight gaps between each response, but the gaps are so small that it is hardly felt by respondents under the normal speed of network connection. With this function, observing
a respondent’s responding sequence and drop out (the point at which he or she stops responding) point can be successfully achieved.

The idea for the program design for this experiment is actually very simple. A frame set based on the HTTP protocol is used to provide these functions. Instead of the real survey page(s), the web server is actually providing two pages inside the frame set, and only one is visible to the respondent. The main page is, of course the questionnaire itself. The other page that provides the interactive submitting behavior is hidden in the back, and occupies no visual space in the survey presentation. Respondents will not notice the existence of this embedded functional page while they are taking the survey. Since the code is less than ten lines in Hypertext Markup Language (HTML) source (listed below*), it is easy to download compared to the questionnaire.

```html
function sendForm(name, value)
{
    document.form1.Rnum.value = name;
    document.form1.Answer.value = value;
    document.form1.Time.value = new Date();
    document.form1.submit();
}
```

*Note: Some corresponding code is also needed at the server side in order to handle data storage. The entire program, including client side handler and server side services is available from the author.

Once the respondent makes a response, the embedded JavaScript code in the
survey page will call the submit function in the hidden page to submit only the changed value back to the server without a flash on the main questionnaire. By doing this, we can successfully collect the responses made by the respondent. And also, the time when the response is made can be recorded into the database for later analysis. This will be a valuable technique for online surveys or experiments if it is able to perform this kind of response observation for us without creating static connection or a delay in program downloading.

Furthermore, in order to achieve this task of distributing different styles of surveys equally to each respondent and processing the data received from respondents, a server side application is also required. Microsoft based ASP page technology (a server side language) and an SQL server database system are used at the web server to help us in automatically executing the online experiment. Basically, we recorded every input from the respondents, even if the respondent had already taken the survey. In this case, an extra row was created for the duplicated response which gave us extra information about the regular online respondents’ behavior. This detailed planning of survey distributing and the data collecting system prior to launching the experiment allows us to distinguish those double responses, and provides us an opportunity for future analysis.

**Survey presenting style**

In order to test the effectiveness of conventional Internet surveys, the electronic questionnaire will be customized into different formats based on several popular Internet survey practices. Regardless of the background design, sound or visual effect, color and data collecting technique behind it, we can roughly categorize conventional Internet surveys into the following three styles.
Style one: the whole questionnaire is presented as a long list in one single page (Long Form). Under this format, respondents will interact with only one page. This depends on the survey length. Usually a survey will not be able to fit inside a regular sized screen. Only a portion of the survey is visible in the browser and respondents need to move the scroll bar (or use the page down button) in order to see the rest of the page.

Style two: the questionnaire is composed of many web pages, and each page contains only one question (Short Form). For this format, one page can usually be displayed in one screen with scrolling. But asking respondents to click the “next” button after they finish the current question is necessary in order to move on. Sometimes, a JavaScript function could be used to advance to the next page automatically. Under the second format, data collection for an incomplete survey (didn’t hit the submit button) is feasible.

Style three: the questionnaire is partitioned into several big chunks (Block Form), and each chunk contains a couple of questions. Each web page will present only one chunk of a question set. Every chunk of the questionnaire is usually the result of some kind of grouping either by type or character of questions such as a whole block of Likert scale questions, a group of demographic questions or a set of sub-questions that share the same survey instruction or heading story.

According to the three commonly used Internet survey styles, three different formats of online questionnaires were made for this study. A “Long Form” that displayed the survey in one page was one (see figure 2 for example). The second design used “Short Form,” and the survey moved to the next question once the respondent clicks the next button (or link) (see figure 3 for example). The third design used the “Block Form” idea where similar questions were grouped together and put into separated pages. In this
format, the length of each chunk (block of page) could be a factor for further consideration. In most cases, we wanted to design the questionnaire in a way that respondents would not have to move the scroll bar in order to view the whole page (see figure 4 for example). It might appear to be less burdensome and easier to answer compared to the second design, because similar types of questions have been put together.

Figure 2. Internet questionnaire in a long list (long form).
Figure 3. Internet questionnaire – One question per page with progress indicator (short form).

2. What is your race or ethnicity?

- AFRICAN AMERICAN/BLACK
- NATIVE AMERICAN
- ASIAN AMERICAN
- EURO-AMERICAN/WHITE
- HISPANIC
- OTHER: [ ]
Figure 4. Internet questionnaire – One block of questions per page without progress indicator (block form).

4. How many minutes did you spend reading any newspapers yesterday? 
   MINUTES

5. How many hours did you spend watching TV yesterday? 
   HOURS

6. About how many friends or family members did you talk with yesterday? 
   NUMBER OF FRIENDS

7. About how many other people did you talk with yesterday? 
   NUMBER OF OTHERS

8. How many times a week do you watch a local television newscast? 
   NUMBER

9. How many times a week do you watch a national television newscast? 
   NUMBER
Figure 5. Internet questionnaire – One block of questions per page with progress indicator (progress block form).

19. Do you own any of the following?

<table>
<thead>
<tr>
<th>YES</th>
<th>NO/DON'T KNOW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Personal computer</td>
</tr>
<tr>
<td></td>
<td>Satellite dish</td>
</tr>
<tr>
<td></td>
<td>VCR</td>
</tr>
<tr>
<td></td>
<td>Digital video disk or DVD</td>
</tr>
<tr>
<td></td>
<td>Video camera</td>
</tr>
<tr>
<td></td>
<td>Compact disk player (CD player)</td>
</tr>
<tr>
<td></td>
<td>Laser disk player (LD player)</td>
</tr>
<tr>
<td></td>
<td>Video game console</td>
</tr>
<tr>
<td></td>
<td>Electronic personal organizer (PDA)</td>
</tr>
<tr>
<td></td>
<td>Electronic pager</td>
</tr>
<tr>
<td></td>
<td>Answering machine</td>
</tr>
<tr>
<td></td>
<td>Cellular telephone</td>
</tr>
<tr>
<td></td>
<td>Fax machine</td>
</tr>
</tbody>
</table>
Basically, the different styles of online survey pages are customized using basic HTML language and the technique we developed for real time data collecting was implanted in each of the corresponding input places.

Besides these three survey styles that are commonly used in the real world, a progress indicator is another feature usually seen. Couper et al. (2000) suggest that the use of a progress indicator can improve the response rate of Internet surveys. Here, adopted one more style (Progress Block Form) to test if the progress indicator really help the response rate by influencing a respondent’s decision to drop out (because the progress indicator appears when they enter the survey). Another group will be tested with a label showing the current page number and total pages based on the 4th presentation design described above (see figure 5 for example).

**Invitation and Survey Delivery**

The method of survey delivery is another factor that might influence a respondent’s intention to respond. Three types of survey invitation methods popular with conventional Internet survey practices will also be tested here. Those methods include the use of e-mail, pop-up window, and an invitation while in the process of some other online activities.

E-mail is the most popular communication method on the Internet. It is also widely used as a survey invitation method in most of the online surveys. The advantage of using an e-mail invitation is that the source of participants is manageable. There are two ways of using e-mail invitation: attach the survey in place, or include a hyperlink that leads the samples to the online survey environment. We used the latter one in this study.

Pop-up window is another popular method to draw Internet users’ attention during
their web surfing activities. It is mostly adopted by a lot of commercial web site or so
called “dot com” companies. One characteristic it provides is that it opens a new window
displaying additional information, and distracts on Internet users’ attention away from
what they were doing to the new window. Since it is also thought to be annoying to some
Internet users, some software companies shipped products to block or disable the popup
action from the popular browsers, and the participants are usually selected for
convenience.

Some other methods are also being developed to effectively draw more
respondents. In this study, the third method we tested was to invite the target participants
while they were performing some online task. It is not as popular as the previous two
methods above, but it does provide the control of a participating source if used with an
appropriate Internet activity procedure design.

**Reward**

Reward has been reported as an important factor in getting more respondents to
traditional survey styles or to the Internet survey we are testing. Different rewards have
been used as incentives in traditional survey research, but few are seen in Internet survey
research. One possible reason is that researchers who choose the Internet as the survey
platform do so because they want to save money. However, if representativeness is the
focus, the quality of response is more important than the quantity. We would like to
follow the suggestion from the survey methodology studies to provide rewards in order to
get a higher response rate instead of just high response numbers based on very big sample
size. Two reward methods were tested here.

**Respondents and Invitation**
Students from Washington State University were the target population in this experiment. The reasons for using student samples were described previously. The reason the Internet is an attractive topic is because of the student accessibility used in classes. The university provides e-mail accounts for each student. Participants in this experiment were grouped by several different categories. Each category was also a factor that might influence either a response or an intention to dropout and therefore were targets we tested in this experiment.

1. **The sampling methods**

   We derived the sample based on two sampling strategies. First we used the probability method. E-mail addresses were randomly selected from the university’s name directory. Since every student in the population is assigned one e-mail account, the random sampling method implemented here is meaningful. 553 were drawn using this method.

   The second strategy we used was to sample by class, which is also used by most studies that use student subjects. Even though some of these studies treated the samples drawn this way as random samples, we preferred to treat ours as a convenience sample. Four different classes were selected from different academic departments to perform different stimuli that will be described later. A total of 570 samples were collected using this sampling method, 177 respondents were collected from a class in the Communication department, 45 from a class in the Decision Science department, about 300 students were from a class in the Business department, and 48 samples were collected from a class in Computer Science. A detailed list and information regarding these grouping methods can be referenced in Table 1.
Table 1. List of grouping categories of samples tested.

<table>
<thead>
<tr>
<th>Group</th>
<th>Sampling type</th>
<th>Sample source</th>
<th>Sample number</th>
<th>Invitation method</th>
<th>Reward method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Random samples</td>
<td>University name directory</td>
<td>553</td>
<td>e-mail</td>
<td>Meal</td>
</tr>
<tr>
<td>2</td>
<td>Communication</td>
<td>Communication</td>
<td>177</td>
<td>e-mail</td>
<td>Extra credits</td>
</tr>
<tr>
<td>3</td>
<td>Convenient samples</td>
<td>Decision Science</td>
<td>45</td>
<td>Pop-up window</td>
<td>Extra credits</td>
</tr>
<tr>
<td>4</td>
<td>Convenient samples</td>
<td>Business</td>
<td>~300</td>
<td>Pop-up window</td>
<td>Meal</td>
</tr>
<tr>
<td>5</td>
<td>Convenient samples</td>
<td>Computer Science</td>
<td>48</td>
<td>Invite in onlinequiz</td>
<td>Extra credits</td>
</tr>
</tbody>
</table>

Implementation of different invitation methods

We divided all the samples into three groups. The e-mail addresses of samples in the first group were collected and the survey invitation was delivered using an automatic e-mail distribution application. Each e-mail contained the invitation message, including information about one of the rewards they were promised, and a customized hyperlink for the online survey encoded with the special identity of each participant. Respondents who clicked on the provided link were directed to the consent form for the survey.

A popup window was used as an invitation tool for the samples drawn from two different classes. A client side function that pops up this invitation window was planted in the class home pages. Students who entered the class web site were given the popup window with the survey invitation on it. Unlike the e-mail invitation which leads respondents to the survey with a key to identify each respondent, the popup window was triggered every time anyone visited the class home page without a predefined identifier. Participating students were asked for their student ID or e-mail address in order to record the extra credit or free meal they earned. The system logged the number of visitors every time the pop up window was triggered, but like most publicly posted online surveys, the
visitor count did not reflect the real number of how many “individuals” had visited the site. For this reason, the real class size was recorded as the real sample size (45 students in the third group and a little bit less than 300 in the fourth group).

The samples in the last group were invited during an online quiz. Besides the questions for the quiz, students were provided a chance to earn extra credit for the quiz they took. In this case, they could only access to the survey while the online quiz was open (students were asked to take the quiz over a week time span of their choosing), once they had submitted the quiz, they could no longer access the survey. A customized hyperlink to the survey including an identifier was encoded in the online quiz, to be able to identify the response of the different respondents. (Please see Table 1)

**Implementation of different reward methods**

As mentioned above, two reward methods were used in this experiment to attract more respondents to the survey. Since students were the majority of our sample and we collected those samples from classes, extra credit was the most convenient and possibly the most useful reward we could use. All students in this group were promised some extra credits for the class they were drawn from.

In order to compare the effectiveness of the different rewards, another reward group was designed and rewards were promised to another group of participants. Respondents from group 1 and 4 were told that they would have a chance to win one of twenty five free meals in a local restaurant (please see Table 1 for number of participants in groups). We provided 25 gift certificates of getting a free meal in the local China Buffet (approximately $10 value) to those who participated. In order to attract more respondents, in the free meal reward group, we adopted the strategy of drawing five free
meal winners each week the day after the reminder e-mail was sent. The selected samples in this group who did not unsubscribe from the mailing list would receive the reminder e-mail with the upcoming drawing for the meal certificates. Utilizing this strategy, the participants selected had some instant motivation upon receiving the reminder e-mail.

**Implementation of presenting different survey styles**

Samples drawn from all the groups were assigned to one of four survey formats by the automatic server-side application. With the help of this application, we could test the influence of the survey format across the groups by different categories. All the transactions, such as identifying the participants (to see whether this participant was a returning respondent or a new respondent), were randomly assigned to a specific group, and then directing the respondent to the desired survey page was done quietly in the back end by the customized server-side application without interruption or delay. In addition, the application used on the server side also guaranteed that if the same respondent came back to the survey after responding the first time that would not be assigned to two different survey styles.

One important issue and often criticized practice is the use of unidentifiable participants. Because of the nature of the Internet, one can visit virtually anywhere as long as he or she knows a valid URL. In this study, personal information such as student ID’s or e-mail addresses were collected before or at the time of the survey in order to identify the duplicated responses from the regular responses. Unlike a posted online survey, every individual has only one chance to enter, much like the traditional survey method. Since e-mail addresses and student ID is unique, we can filter duplicated e-mail using the computer program before launching the survey. The classes we selected for the
The last four groups were from the fields of social science, business and engineering, so the chance that one student would be selected twice is not likely even though they were not selected randomly like those in group one. Once the experiment stopped, and the extra credit reward was administered, the personal information was destroyed. A meaningless ID was used to replace the original personal identification for statistical requirements afterwards and for confidentiality.

For the two groups who were received an e-mail invitation to enter the survey, we also provided a tool to unsubscribe them from the reminder list. This is used popularly in conventional e-mail advertising or newsgroups. If they didn’t unsubscribe from the list by themselves, the e-mail distribution system would send an e-mail reminder every week that the survey was active if the selected respondent had not yet responded to the survey.

2. **Experimental Procedure**

The experimental procedure started after the respondent agreed to the consent form. At the beginning of the experiment, the respondents’ identifying information, (either e-mail address or student ID), was verified by the server side application. If it was the first time the respondent entered the survey, one of the four survey styles was randomly and evenly assigned to the respondent. On the other hand, if respondents re-enter the survey by any chance, the style they were originally assigned to was retrieved from the database and presented to them. While taking the survey, like previously described, every time a respondent makes a response, the client side JavaScript function will automatically send the information, including the question number and the value they entered back to the server to be recorded. So even if a respondent decided not to complete the survey, we would still have all of the information they had entered. When recording
each response, our server side application also put a time stamp on the response.

The experiment lasted for three weeks, including one week of Thanksgiving break. Unless participants either responded or unsubscribed to the mailing list, an e-mail reminder was sent to them each week containing the same customized survey link.

After the experiment ended, the survey page was taken out of the public web space, and the various rewards were distributed. Once the reward procedure was done, all the personal identification information was deleted from the database, so we knew the respondent ID instead of their personal information.
CHAPTER FIVE
RESULTS AND ANALYSIS

Of the three different survey distribution methods used in this study, the invitation popup window group couldn’t provide an exact number of the amount of invitations delivered. The number of times the popup window was triggered was not the actual value of the survey information that was sent. The same person who visited the class home page ten times would have triggered the program to bring up the survey invitation window ten times. This is also the reason that the real response rate of a posted Internet survey is difficult to evaluate. Since we asked the instructors of the participating classes to announce the online survey event, every student of the sample in those classes already received the invitation message prior to visiting the class site. We can assume that all of the students in the selected classes for the popup window invitation group were potential participants. If we count the number of students in those classes and the invitation e-mails we sent plus the students who took the online quiz, the total number of potential respondents was 1123. Regardless of the invitation method, 308 of those potential respondents actually attempted to respond, or about 27.4% (Please see Table 2). This number is close to the high average in most of the online research reported. Considering that we also included a mechanism for the respondents who received the invitation e-mail to unsubscribe themselves from future reminder e-mails, those who unsubscribed could also be treated as “responded” respondents. These respondents received the information we sent and did respond to one of the functions provided. Actually, only six of the 730 respondents who received our invitation e-mail used this tool to unsubscribe, a mere 0.8%, or 1% of the samples drawn from the school name directory were given a free meal.
instead of the extra credit as a promised reward. What this shows is that even though providing a way for recipients to exclude themselves from future e-mails from the sender, (a requirement for not being defined as a spam), few e-mail users actually utilize it.

**Demographics**

From the valid responses, 160 of them were male (55.6%), and 128 were female (44.4%). The majority of the participants identified themselves as Euro-American/White (232, 82.9%), 7 were African American/Black (2.5%), 3 were Native American (1.1%), 14 were Asian American (5%), and 6 were Hispanic (2.2%). The respondents’ age range was from 18 to 48 with the average being 22. Interestingly, 117 (42%) of them identified themselves as heavy Internet users, 135 (48.6%) were medium users, and only 23 (8.3%) of them see themselves as light Internet users. Also, 218 (70%) of the respondents reported having Internet access both at home and at school. 266 of those respondents described themselves as somewhat knowledgeable about the Internet, which was almost 90% of all respondents. Results revealed that the original assumption about the common Internet access and average computer literacy about the target population was pretty much correct. It also indicated that we tested a homogeneous pool of respondents in terms of Internet usage.

**Survey completeness**

Because the main interest here is testing the drop out rate, we focused on the 308 respondents who actually responded to the survey. Here, we looked at the data we collected from the three different aspects of the different grouping strategies, by *invitation method*, by *rewarding method* and by *survey presentation style*. We did the analysis by first treating all responses as one big group, and then did the comparison of
the grouping of interest. For example, when we analyzed the response rate and the
dropout rate by different reward groups, we disregarded the fact that the responses were
from 5 different groups, and separated them into two groups based on the different
rewards promised.

Our data shows that of those 308 respondents who actually attempted to respond,
5 of them stopped without answering any questions, 28 of them dropped out from the
survey before hitting the submit button. The incomplete rate was about 10.71% including
both cases (Please see Table 2 for a detailed number). If we had conducted the survey in
the traditional way, collecting respondent answers after they have hit the final submit
button, then those 33 incomplete surveys would have been counted as non-responses. In
this comparison, the dropout rate is not high enough to stand out from the total
non-response rate. Hence, the first hypothesis, which predicted that the dropout rate and
the non-response rate would be different using the row data without any grouping strategy,
was not supported due to the relatively high non-response rate. Also, from the gender
information collected from responded samples, we found no difference in the drop out
rate in terms of gender. On the other hand, of the 275 respondents who completed the
survey (and finished the survey by clicking the submit button), we found that there were
12 (3.6%) of them that left over 10 question items unanswered out of a total of 58
question items, (please reference the appendix for the question item list). The results of
the response rate and the dropout rate by different groups are described below.
Table 2. Response results group by invitation

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inv. sent</td>
<td>Unsubscribed</td>
<td>Non-response</td>
<td>Responded</td>
<td>Stop before survey begin</td>
<td>Drop out</td>
<td>Double respond</td>
<td>Partial respond</td>
</tr>
<tr>
<td>Extra Credit</td>
<td>177</td>
<td>0(0)</td>
<td>36(20.3)</td>
<td>141(79.7)</td>
<td>1(0.7)</td>
<td>12(8.5/6.8)</td>
<td>10(7.1)</td>
<td>3(2.1)</td>
</tr>
<tr>
<td>Free Meal</td>
<td>553</td>
<td>6(1.0)</td>
<td>482(87.2)</td>
<td>65(11.8)</td>
<td>4(6.2)</td>
<td>8(12.3/1.5)</td>
<td>1(1.5)</td>
<td>3(4.6)</td>
</tr>
<tr>
<td>Subtotal</td>
<td>730</td>
<td>6(0.8)</td>
<td>518(71.0)</td>
<td>206(28.2)</td>
<td>5(2.4)</td>
<td>20(9.7/2.7)</td>
<td>11(5.3)</td>
<td>6(2.9)</td>
</tr>
<tr>
<td>Group 2. Invite using popup window triggered by visiting the class webpage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra Credit</td>
<td>319(45)</td>
<td>-</td>
<td>-</td>
<td>27</td>
<td>0</td>
<td>1(3.7)</td>
<td>2(7.4)</td>
<td>3(11.1)</td>
</tr>
<tr>
<td>Free Meal</td>
<td>878(~300)</td>
<td>-</td>
<td>-</td>
<td>28</td>
<td>0</td>
<td>4(14.3)</td>
<td>3(10.7)</td>
<td>0</td>
</tr>
<tr>
<td>Subtotal</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>55</td>
<td>0</td>
<td>5(9.1)</td>
<td>5(9.1)</td>
<td>3(5.5)</td>
</tr>
<tr>
<td>Group 3. Invite in online quiz (as the format of one question in the quiz)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra Credit</td>
<td>48</td>
<td>-</td>
<td>1(2.1)</td>
<td>47(97.9)</td>
<td>0</td>
<td>3(6.4/3.3)</td>
<td>0</td>
<td>3(6.4)</td>
</tr>
<tr>
<td>Grand Total</td>
<td>1123</td>
<td>6(0.5)</td>
<td>308</td>
<td>5(1.6)</td>
<td>28(9.1)</td>
<td>16(5.2)</td>
<td>12(3.9)</td>
<td></td>
</tr>
</tbody>
</table>

1 One of the respondents who unsubscribed also responded to the survey.
1. The invitation method

Of the 730 invitation e-mails sent, 518 (70.1%) of them didn’t respond to the survey (they neither responded nor unsubscribed). Among the 212 who did respond, 6 unsubscribed, and another 20 (9.7% of 206 response respondents) dropped out without submitting the survey. If we add the 20 dropouts to the 518 non-responses, we end up comparing 518 with 538 (non-responses including dropouts versus non-responses only), the difference was about 4%. However, the data leak which counted the dropouts as a part of the non-responses (without catching the dropouts) was not very serious. The first hypothesis was not supported in this group.

For the popup window group, we used the class size as the number of potential participants. Only 55 (15%) out of nearly 350 potential participants responded to the survey. And within the 55 respondents, 5 (9.1%) of them dropped out before they submitted. Because of the nature of using popup windows, we were not able to predict how many “individuals” triggered the Internet survey. To calculate the non-response rate using either the number of times the popup windows was triggered or the actual class size would likely introduce more deviation. If we consider that every student in the classes we drew samples from were aware of the survey information, then the drop out amount was not comparable to those who didn’t take action at all. The first hypothesis was not supported either.

Finally, for the group that participated in the online experiment during an online quiz, almost all of them responded (47 out of 48, a 98% response rate), and 3 (6.4%) of them dropped out in the middle (Table 2). Here, we could see that the number of dropouts was even more than the number of non-responses. Comparing the non-response rate that
was generated with and without the ability to capture the dropout events, we could see that “one” non-response (without dropouts) is definitely different from “four” (with dropouts). Because in this case, the factors that influenced the respondents to leave the survey was more important than the factors that failed to attract the samples to the survey. By using the old method without capturing the dropouts, we would be distracted and lose the real factors that cause incomplete responses. In this group, the first hypothesis was supported.

From the data generated from the three different invitation methods, we can see that different survey delivery methods do have different effects on survey response rates. The method that invited the target participants in the middle of some relevant online tasks appears to be the best for survey response (highest response rate). Hence, H2b was supported for effect of different invitation method on response rates. However, the difference of the drop out rate between the three groups was not big enough to support the conclusion that the invitation method was an important factor in the drop out rate. H2c was not supported here.

Since the invitation method provides the first impression about the survey, it was reasonable that it has more influence in attracting samples to the survey instead of retaining them. We can conclude that the invitation methods of Internet survey deploying plays a more important role in the first step of a respondent’s decision-making procedure.

2. The Reward method

After seeing the number in the invitation methods, and scrutinizing the response by the type of reward (extra credit or free meal), we found the results to be quite different (see Table 3 for detail). Of the 270 participants who were given a chance to win extra
credit points in class, 215 of them responded, an outstanding 79.6% response rate. However, of the 853 potential participants who were given a chance to win a free meal at the local restaurant, only 93 of them responded, a little bit over a 10% response rate. Comparing the drop out rate between these two reward systems, 17 from the extra credit system didn’t complete the survey (7.9% in extra credit system), and 16 of these from the free meal system (17.2% in free meal system).

If we look at those numbers from the traditional point of view, which treats all dropouts as non-responses, we find something different (please reference to Table 3, column F, G and H). For the group that was promised extra credit, there were a total of 72 that didn’t reach the final submit step (didn’t push the submit button, treated as non-responses in the traditional point of view). 55 of the 72 didn’t make the attempt to respond at all (treated as non-response in this study), and 17 dropped out in the middle of the survey (treated as dropouts in this study, along with the gap between the two ways of seeing the non-response rate). The difference between the two analysis methods were the 17 dropouts would have been included in the non-responses if we hadn’t monitored the dropout behavior in our study. Almost 25% of the information would have leaked if we were unable to extract the dropouts from the respondents who were not interested in giving the survey a look at all! Technically, this 25% is the effort we can make in survey design (instead of message delivery) to keep respondents more interested for longer in the survey to the point of final submission. So H1 was supported using the extra credit reward. Compared to the group that was promised free meals, a total of 776 didn’t submit the survey, 760 of them didn’t respond at all, so only 16 (2%) of them dropped out. H1 was not supported for this group.
Table 3. Response results group by rewarding method

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Responded (% of A)</td>
<td></td>
<td></td>
<td>Unsubmitted</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(% of A) [B + E]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total invited</td>
<td></td>
<td></td>
<td>Submitted (% of C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Drop-out (% of C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Non-response (% of F)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Drop-out (% of F)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Partial responded</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(% of D)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra Credit</td>
<td>270</td>
<td>55 (20.0%)</td>
<td>215 (80.0%)</td>
<td></td>
<td>72 (27.0%)</td>
<td></td>
<td></td>
<td>9 (4.5%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>198 (92.0%)</td>
<td>17 (7.9%)</td>
<td>55 (76.0%)</td>
<td>17 (24.0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free Meal</td>
<td>853</td>
<td>760 (89.0%)</td>
<td>93 (11%)</td>
<td></td>
<td>776</td>
<td></td>
<td></td>
<td>3 (3.9%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>77 (83.0%)</td>
<td>16 (17.2%)</td>
<td>760 (98.0%)</td>
<td>16 (2.0%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
From the results in Table 3, we can see that extra credit drew an outstanding 80% response rate, while free meal gained only 11%. We can say that H2b was supported for different incentives. And the drop out rate comparison between the two methods proved there was difference. Thus H2c was also accepted. Apparently, the incentive used influenced the response rate than drop out rate. By analyzing the influence of incentives used, it provided the information about what kind of reward will collect better results. It showed that a more attractive reward does provide invited recipients with more motivation to start the survey. It was also more important in the first step of a respondent’s decision-making procedure.

3. The survey presentation style

In terms of the survey presentation style, we first took a look at how effective our program evenly distributed different survey styles to each respondent regardless of the invitation or reward methods. What we found was that the server side application we designed successfully distributed the survey with different styles (long form, short form, block form and progress indicator form) evenly regardless of what group the respondent was drawn from or which invitation method lead them to the survey. We could see this from the total number of respondents in each group in Table 3.

Under this grouping strategy, each respondent could have been drawn with any invitation method or any reward incentive, and would have no way of knowing what group the samples who didn’t take any action were distributed in. In addition, respondents had no idea at all about the survey presentation format prior to entering the survey. When they decided to enter the survey, they were no longer non-respondents. We could then safely assume that the survey presentation style influenced the respondents’
decision-making only in terms of completing or dropping out, not in terms of whether to take the action to answer. For this reason, we were unable to compare the non-response rate with the dropout rate if the survey presentation style was not one of the factors that motivated them to attempt to respond. This hypothesis will be discussed later.

Here we focused our analysis on how the survey presentation styles influenced the decision to dropout (during step 3 of Figure 1). According to our data, the dropout rates between these four groups were different, but the difference was not big enough for us to perform further analysis. The number of dropouts was also too small to provide a basis for meaningful statistical analysis. We could see that the style with all of the question items on one page and the style with one question item per page had the highest dropout rates (11.84%). The style that had similar question items on one page along with the progress indicator received the lowest dropout rate of 5.3%. Detailed data can be referenced in Table 4.

Apparently, the dropout rates that were collected from the groups based on the survey presentation style couldn’t provide enough evidence about the differences between the four different styles. Even though the trend of the block form with the progress indicator kept respondents in the survey longer, and retained more complete responses than the other styles. Without solid statistical analysis, we cannot say that hypothesis H2a was supported.
Table 4. Response results group by survey format

<table>
<thead>
<tr>
<th>Format</th>
<th>Total (Who agree on the consent)</th>
<th>Stop (Didn’t answer any question)</th>
<th>Enter survey</th>
<th>Drop out</th>
<th>Survey submitted</th>
<th>Average Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Completed</td>
<td>Partial completed</td>
</tr>
<tr>
<td>Format 1</td>
<td>78</td>
<td>2</td>
<td>76</td>
<td>9 (11.8%)</td>
<td>67 (88.2%)</td>
<td>64</td>
</tr>
<tr>
<td>Format 2</td>
<td>77</td>
<td>1</td>
<td>76</td>
<td>9 (11.8%)</td>
<td>67 (88.2%)</td>
<td>66</td>
</tr>
<tr>
<td>Format 3</td>
<td>76</td>
<td>1</td>
<td>75</td>
<td>6 (8.0%)</td>
<td>69 (92.0%)</td>
<td>63</td>
</tr>
<tr>
<td>Format 4</td>
<td>77</td>
<td>1</td>
<td>76</td>
<td>4 (5.3%)</td>
<td>72 (94.8%)</td>
<td>70</td>
</tr>
</tbody>
</table>

1. A long questionnaire form with all the question items (long form).
2. One question item per page for the whole questionnaire. This form has total of 58 pages. Also have progress indicator for this format (short form).
3. One block of questions per page that includes similar item of questions inside each page. It has a total of 8 pages, and doesn’t have the progress indicator (block form).
4. The same as format 3. A progress indicator differentiates it from the format 3 (progress block form).
Time spent on the survey

Something that went very well was the program we designed to collect the response behavior information including each move a respondent made while their participating in the online survey. By collecting this information, we were able to know the way each person responded even when they gave up in the middle or failed to submit the survey. In addition, we collected information about when they answered each question. Hence, the time they spent on each question or the whole survey could be calculated. Our record shows that the average time it took to finish this online survey regardless of the presentation style was about 420 seconds (7 minutes). But when we looked more closely, we found that the time it took to finish the survey depended on which survey presentation style the respondent was assigned to.

It was not surprising to find that the survey that was presented as a long list of questions took shortest time (347.8 seconds) to finish, and the one that separated every question item on different web pages took the longest time (463.4 seconds) to finish (see Table 5 for detail data). This was because the respondents needed to click two times to answer one question item instead of the one click in the first format. From the data of time spent on each question (see Table 6 for detail), we could see that the questions in format 1 took less time to answer than the other 3 presentation styles. Figure 6 and figure 7 are graphs that show the difference between groups. In figure 6, we could see that some questions required more time to answer across all four of the survey styles we presented. For example question item number 6, which asks how many friends or family members they talked to yesterday, appeared to be the highest across all four groups. On the other hand, some high points were only in format 3 and 4 and appeared to be the first question on a new page, such as question item 10 and 35. These format specific high points were
not as obvious in format one and two.

By using the time log on each question answered, we were also able to identify where the respondent dropped out of the survey. By grouping the time spent on each question by the completed survey response (Table 7 and Figure 8), we could see that all of the respondents who dropped out of the survey dropped out before question item 43. Actually, upon further inspection of the drop out point of each of these respondents; about half dropped out at the 10th question. Some of them stayed in the survey for a while longer, but the total time spent on the survey was approximately 300 seconds. Overall, the longer they spent on the survey, the more questions they answered.

However, the drop out points varied from respondent to respondent, but it appears that most of the respondents who dropped out of the experiment answered question item 9, which was at the end of the second page. In questions 1 to 9, we were asking simple facts about basic demographic information and recent media usage (referenced in the appendix). From question 10, we began to ask respondents their feeling about some issues, which was not that straight forward and usually, requires more time and cognitive energy. About one third of the respondents who dropped out actually stopped after they answered the 10th question item. This shows that the difficulty of the task they perceived influenced their intention to answer. A dropout point versus a dropout time figure (Figure. 9) shows the dropout point and time distribution more clearly.
Table 5. Accumulate time spent on survey by survey presentation style. (Express in second).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>All response</td>
<td>10.1</td>
<td>14.9</td>
<td>21.6</td>
<td>33.1</td>
<td>45.2</td>
<td>66.3</td>
<td>77.8</td>
<td>87.2</td>
<td>96.3</td>
<td>124.7</td>
<td>131.0</td>
<td>137.9</td>
<td>157.7</td>
<td>164.6</td>
</tr>
<tr>
<td>Format 1</td>
<td>11.8</td>
<td>14.7</td>
<td>22.9</td>
<td>30.1</td>
<td>40.7</td>
<td>56.0</td>
<td>63.9</td>
<td>75.1</td>
<td>86.5</td>
<td>99.4</td>
<td>106.1</td>
<td>120.6</td>
<td>128.5</td>
<td>133.0</td>
</tr>
<tr>
<td>Format 2</td>
<td>5.6</td>
<td>15.8</td>
<td>22.3</td>
<td>34.0</td>
<td>48.6</td>
<td>71.2</td>
<td>85.5</td>
<td>95.7</td>
<td>102.9</td>
<td>120.0</td>
<td>128.0</td>
<td>136.0</td>
<td>159.4</td>
<td>169.2</td>
</tr>
<tr>
<td>Format 3</td>
<td>12.8</td>
<td>15.4</td>
<td>22.1</td>
<td>35.8</td>
<td>47.9</td>
<td>72.0</td>
<td>84.5</td>
<td>92.9</td>
<td>105.1</td>
<td>139.4</td>
<td>144.4</td>
<td>146.9</td>
<td>170.6</td>
<td>178.5</td>
</tr>
<tr>
<td>Format 4</td>
<td>10.6</td>
<td>13.8</td>
<td>19.1</td>
<td>32.2</td>
<td>43.6</td>
<td>66.0</td>
<td>77.2</td>
<td>85.2</td>
<td>91.1</td>
<td>137.5</td>
<td>143.1</td>
<td>145.9</td>
<td>169.8</td>
<td>175.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
</tr>
</thead>
<tbody>
<tr>
<td>All response</td>
<td>170.2</td>
<td>176.6</td>
<td>182.9</td>
<td>187.4</td>
<td>192.1</td>
<td>199.3</td>
<td>204.5</td>
<td>210.9</td>
<td>218.9</td>
<td>226.6</td>
<td>232.1</td>
<td>238.4</td>
<td>244.0</td>
<td>249.7</td>
</tr>
<tr>
<td>Format 1</td>
<td>139.7</td>
<td>145.1</td>
<td>150.5</td>
<td>154.6</td>
<td>159.4</td>
<td>162.9</td>
<td>168.8</td>
<td>174.1</td>
<td>181.1</td>
<td>188.5</td>
<td>192.3</td>
<td>199.8</td>
<td>204.7</td>
<td>209.1</td>
</tr>
<tr>
<td>Format 2</td>
<td>176.9</td>
<td>183.0</td>
<td>189.6</td>
<td>195.6</td>
<td>195.6</td>
<td>212.9</td>
<td>217.8</td>
<td>226.6</td>
<td>231.9</td>
<td>241.3</td>
<td>247.2</td>
<td>253.3</td>
<td>259.6</td>
<td>267.6</td>
</tr>
<tr>
<td>Format 3</td>
<td>180.6</td>
<td>190.8</td>
<td>198.8</td>
<td>201.8</td>
<td>206.6</td>
<td>211.1</td>
<td>215.9</td>
<td>221.8</td>
<td>232.7</td>
<td>240.4</td>
<td>246.6</td>
<td>252.8</td>
<td>257.9</td>
<td>262.6</td>
</tr>
<tr>
<td>Format 4</td>
<td>180.9</td>
<td>185.3</td>
<td>190.6</td>
<td>195.2</td>
<td>199.7</td>
<td>203.4</td>
<td>208.5</td>
<td>214.2</td>
<td>222.8</td>
<td>229.4</td>
<td>235.2</td>
<td>240.9</td>
<td>247.0</td>
<td>252.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>29</th>
<th>30</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
<th>41</th>
<th>42</th>
</tr>
</thead>
<tbody>
<tr>
<td>All response</td>
<td>255.6</td>
<td>262.6</td>
<td>267.5</td>
<td>272.2</td>
<td>278.0</td>
<td>283.2</td>
<td>297.1</td>
<td>304.3</td>
<td>308.4</td>
<td>311.5</td>
<td>313.6</td>
<td>315.1</td>
<td>352.0</td>
<td>360.0</td>
</tr>
<tr>
<td>Format 1</td>
<td>214.0</td>
<td>219.0</td>
<td>223.4</td>
<td>227.9</td>
<td>234.7</td>
<td>240.9</td>
<td>245.8</td>
<td>251.0</td>
<td>254.7</td>
<td>257.8</td>
<td>260.5</td>
<td>263.0</td>
<td>306.2</td>
<td>308.2</td>
</tr>
<tr>
<td>Format 2</td>
<td>272.7</td>
<td>279.5</td>
<td>285.4</td>
<td>291.0</td>
<td>295.9</td>
<td>303.2</td>
<td>311.9</td>
<td>320.5</td>
<td>327.3</td>
<td>331.8</td>
<td>333.5</td>
<td>336.5</td>
<td>367.4</td>
<td>375.1</td>
</tr>
<tr>
<td>Format 3</td>
<td>270.9</td>
<td>283.1</td>
<td>288.4</td>
<td>292.4</td>
<td>298.7</td>
<td>304.4</td>
<td>334.3</td>
<td>341.2</td>
<td>346.0</td>
<td>348.1</td>
<td>350.1</td>
<td>352.1</td>
<td>392.7</td>
<td>404.5</td>
</tr>
<tr>
<td>Format 4</td>
<td>258.0</td>
<td>262.1</td>
<td>266.0</td>
<td>270.9</td>
<td>276.1</td>
<td>278.1</td>
<td>290.7</td>
<td>298.7</td>
<td>299.5</td>
<td>302.1</td>
<td>304.2</td>
<td>303.0</td>
<td>336.6</td>
<td>347.2</td>
</tr>
<tr>
<td></td>
<td>43</td>
<td>44</td>
<td>45</td>
<td>46</td>
<td>47</td>
<td>48</td>
<td>49</td>
<td>50</td>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
<td>56</td>
</tr>
<tr>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>All response</td>
<td>363.0</td>
<td>365.5</td>
<td>368.2</td>
<td>371.0</td>
<td>380.8</td>
<td>384.2</td>
<td>388.6</td>
<td>391.0</td>
<td>393.5</td>
<td>396.7</td>
<td>399.7</td>
<td>403.5</td>
<td>411.0</td>
<td>414.1</td>
</tr>
<tr>
<td>Format 1</td>
<td>310.9</td>
<td>313.5</td>
<td>315.6</td>
<td>317.9</td>
<td>318.6</td>
<td>321.9</td>
<td>324.6</td>
<td>326.3</td>
<td>328.1</td>
<td>329.5</td>
<td>331.3</td>
<td>337.4</td>
<td>342.2</td>
<td>343.8</td>
</tr>
<tr>
<td>Format 2</td>
<td>380.7</td>
<td>384.9</td>
<td>389.4</td>
<td>393.9</td>
<td>398.1</td>
<td>403.0</td>
<td>411.9</td>
<td>417.8</td>
<td>422.6</td>
<td>429.8</td>
<td>437.0</td>
<td>442.2</td>
<td>446.6</td>
<td>453.1</td>
</tr>
<tr>
<td>Format 3</td>
<td>406.2</td>
<td>408.1</td>
<td>410.3</td>
<td>412.7</td>
<td>420.0</td>
<td>421.9</td>
<td>424.2</td>
<td>425.9</td>
<td>427.7</td>
<td>429.3</td>
<td>430.7</td>
<td>433.0</td>
<td>443.6</td>
<td>446.4</td>
</tr>
<tr>
<td>Format 4</td>
<td>348.8</td>
<td>350.3</td>
<td>352.3</td>
<td>354.2</td>
<td>380.1</td>
<td>383.7</td>
<td>387.1</td>
<td>387.6</td>
<td>389.2</td>
<td>391.3</td>
<td>392.7</td>
<td>394.7</td>
<td>404.8</td>
<td>406.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>57</th>
<th>58</th>
</tr>
</thead>
<tbody>
<tr>
<td>All response</td>
<td>416.5</td>
<td>419.7</td>
</tr>
<tr>
<td>Format 1</td>
<td>345.6</td>
<td>347.8</td>
</tr>
<tr>
<td>Format 2</td>
<td>458.2</td>
<td>463.4</td>
</tr>
<tr>
<td>Format 3</td>
<td>448.2</td>
<td>450.2</td>
</tr>
<tr>
<td>Format 4</td>
<td>407.1</td>
<td>410.6</td>
</tr>
</tbody>
</table>

Note: Format 1 through 4 as described in Table 4.
Table 6. Time spent on each question item by survey presentation style. (Express in second).

<table>
<thead>
<tr>
<th>Format</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Format 1</td>
<td>11.8</td>
<td>2.9</td>
<td>8.2</td>
<td>7.3</td>
<td>10.6</td>
<td>15.3</td>
<td>7.9</td>
<td>11.2</td>
<td>11.4</td>
<td>12.9</td>
<td>6.7</td>
<td>14.5</td>
<td>7.8</td>
<td>4.5</td>
<td>6.7</td>
</tr>
<tr>
<td>Format 2</td>
<td>5.6</td>
<td>10.2</td>
<td>6.5</td>
<td>11.7</td>
<td>14.6</td>
<td>22.6</td>
<td>14.3</td>
<td>10.2</td>
<td>7.2</td>
<td>17.1</td>
<td>8.0</td>
<td>8.1</td>
<td>23.4</td>
<td>9.8</td>
<td>7.7</td>
</tr>
<tr>
<td>Format 3</td>
<td>12.8</td>
<td>2.6</td>
<td>6.8</td>
<td>13.7</td>
<td>12.1</td>
<td>24.1</td>
<td>12.5</td>
<td>8.4</td>
<td>12.2</td>
<td>34.4</td>
<td>5.0</td>
<td>2.4</td>
<td>23.7</td>
<td>8.0</td>
<td>2.1</td>
</tr>
<tr>
<td>Format 4</td>
<td>10.6</td>
<td>3.2</td>
<td>5.2</td>
<td>13.1</td>
<td>11.4</td>
<td>22.2</td>
<td>11.0</td>
<td>8.0</td>
<td>6.0</td>
<td>46.0</td>
<td>5.6</td>
<td>2.8</td>
<td>23.9</td>
<td>5.5</td>
<td>5.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Format</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Format 1</td>
<td>5.4</td>
<td>5.4</td>
<td>4.1</td>
<td>4.8</td>
<td>3.6</td>
<td>5.8</td>
<td>5.3</td>
<td>7.0</td>
<td>7.4</td>
<td>3.8</td>
<td>7.5</td>
<td>4.9</td>
<td>4.4</td>
<td>4.9</td>
<td>5.0</td>
</tr>
<tr>
<td>Format 2</td>
<td>6.1</td>
<td>6.6</td>
<td>6.1</td>
<td>--</td>
<td>17.2</td>
<td>4.9</td>
<td>8.8</td>
<td>5.3</td>
<td>9.5</td>
<td>5.9</td>
<td>6.1</td>
<td>6.3</td>
<td>8.1</td>
<td>5.1</td>
<td>6.8</td>
</tr>
<tr>
<td>Format 3</td>
<td>10.2</td>
<td>8.1</td>
<td>3.0</td>
<td>4.8</td>
<td>4.5</td>
<td>4.8</td>
<td>5.9</td>
<td>11.0</td>
<td>7.7</td>
<td>6.2</td>
<td>6.2</td>
<td>5.1</td>
<td>4.7</td>
<td>8.2</td>
<td>12.3</td>
</tr>
<tr>
<td>Format 4</td>
<td>4.4</td>
<td>5.3</td>
<td>4.8</td>
<td>4.3</td>
<td>3.8</td>
<td>5.1</td>
<td>5.6</td>
<td>8.6</td>
<td>6.5</td>
<td>5.8</td>
<td>5.7</td>
<td>6.0</td>
<td>5.3</td>
<td>5.6</td>
<td>4.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Format</th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
<th>41</th>
<th>42</th>
<th>43</th>
<th>44</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>Format 1</td>
<td>4.4</td>
<td>4.4</td>
<td>6.8</td>
<td>6.2</td>
<td>4.9</td>
<td>5.2</td>
<td>3.7</td>
<td>3.1</td>
<td>2.7</td>
<td>2.5</td>
<td>43.2</td>
<td>1.9</td>
<td>2.8</td>
<td>2.5</td>
<td>2.2</td>
</tr>
<tr>
<td>Format 2</td>
<td>5.9</td>
<td>5.6</td>
<td>4.9</td>
<td>7.3</td>
<td>8.7</td>
<td>8.6</td>
<td>6.8</td>
<td>4.5</td>
<td>1.7</td>
<td>3.0</td>
<td>30.9</td>
<td>7.7</td>
<td>5.6</td>
<td>4.2</td>
<td>4.5</td>
</tr>
<tr>
<td>Format 3</td>
<td>5.3</td>
<td>4.1</td>
<td>6.3</td>
<td>5.7</td>
<td>29.9</td>
<td>6.9</td>
<td>4.7</td>
<td>2.1</td>
<td>2.1</td>
<td>2.0</td>
<td>40.5</td>
<td>11.9</td>
<td>1.7</td>
<td>1.9</td>
<td>2.2</td>
</tr>
<tr>
<td>Format 4</td>
<td>4.0</td>
<td>4.9</td>
<td>5.2</td>
<td>2.0</td>
<td>12.6</td>
<td>8.0</td>
<td>0.8</td>
<td>2.6</td>
<td>2.0</td>
<td>-1.2</td>
<td>33.7</td>
<td>10.5</td>
<td>1.6</td>
<td>1.6</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>47</td>
<td>48</td>
<td>49</td>
<td>50</td>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
<td>56</td>
<td>57</td>
<td>58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Format 1</td>
<td>2.3</td>
<td>0.7</td>
<td>3.3</td>
<td>2.7</td>
<td>1.7</td>
<td>1.8</td>
<td>1.4</td>
<td>1.9</td>
<td>6.1</td>
<td>4.7</td>
<td>1.6</td>
<td>1.8</td>
<td>2.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Format 2</td>
<td>4.5</td>
<td>4.1</td>
<td>4.9</td>
<td>8.9</td>
<td>5.9</td>
<td>4.8</td>
<td>7.2</td>
<td>7.2</td>
<td>5.2</td>
<td>4.4</td>
<td>6.6</td>
<td>5.1</td>
<td>5.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Format 3</td>
<td>2.4</td>
<td>7.3</td>
<td>1.8</td>
<td>2.3</td>
<td>1.7</td>
<td>1.7</td>
<td>1.7</td>
<td>1.4</td>
<td>2.3</td>
<td>10.7</td>
<td>2.8</td>
<td>1.8</td>
<td>2.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Format 4</td>
<td>1.9</td>
<td>25.6</td>
<td>3.6</td>
<td>3.3</td>
<td>0.6</td>
<td>1.6</td>
<td>2.0</td>
<td>1.4</td>
<td>2.0</td>
<td>10.2</td>
<td>1.6</td>
<td>0.7</td>
<td>3.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7. Time spent on each question item by survey completeness. (Express in second).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>10.7</td>
<td>4.8</td>
<td>7.6</td>
<td>12.9</td>
<td>11.8</td>
<td>21.0</td>
<td>11.4</td>
<td>9.5</td>
<td>9.8</td>
<td>27.7</td>
<td>6.4</td>
<td>6.8</td>
<td>19.6</td>
<td>6.9</td>
<td>5.6</td>
</tr>
<tr>
<td>Valid</td>
<td>10.2</td>
<td>4.8</td>
<td>6.6</td>
<td>11.5</td>
<td>12.2</td>
<td>21.1</td>
<td>11.5</td>
<td>9.4</td>
<td>9.1</td>
<td>28.3</td>
<td>6.3</td>
<td>6.9</td>
<td>19.8</td>
<td>6.9</td>
<td>5.6</td>
</tr>
<tr>
<td>Partial C.</td>
<td>13.4</td>
<td>2.5</td>
<td>4.7</td>
<td>19.2</td>
<td>5.4</td>
<td>31.4</td>
<td>21.1</td>
<td>6.6</td>
<td>6.5</td>
<td>24.8</td>
<td>5.3</td>
<td>5.8</td>
<td>19.5</td>
<td>5.4</td>
<td>5.4</td>
</tr>
<tr>
<td>Incomplete</td>
<td>7.3</td>
<td>5.8</td>
<td>15.7</td>
<td>11.1</td>
<td>10.2</td>
<td>13.6</td>
<td>7.2</td>
<td>11.8</td>
<td>6.7</td>
<td>12.3</td>
<td>11.8</td>
<td>5.3</td>
<td>9.6</td>
<td>7.6</td>
<td>4.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>6.3</td>
<td>6.3</td>
<td>5.2</td>
<td>4.2</td>
<td>7.6</td>
<td>5.1</td>
<td>6.3</td>
<td>7.9</td>
<td>7.7</td>
<td>5.5</td>
<td>6.3</td>
<td>5.6</td>
<td>5.8</td>
<td>5.9</td>
<td>6.9</td>
</tr>
<tr>
<td>Valid</td>
<td>6.4</td>
<td>6.3</td>
<td>4.5</td>
<td>4.6</td>
<td>7.3</td>
<td>5.2</td>
<td>6.4</td>
<td>7.9</td>
<td>7.8</td>
<td>5.5</td>
<td>6.3</td>
<td>5.6</td>
<td>5.7</td>
<td>5.9</td>
<td>7.0</td>
</tr>
<tr>
<td>Partial C.</td>
<td>5.2</td>
<td>7.2</td>
<td>14.7</td>
<td>-6.3</td>
<td>5.9</td>
<td>4.7</td>
<td>4.3</td>
<td>6.6</td>
<td>7.3</td>
<td>6.7</td>
<td>6.2</td>
<td>6.7</td>
<td>6.6</td>
<td>4.2</td>
<td>4.1</td>
</tr>
<tr>
<td>Incomplete</td>
<td>4.0</td>
<td>4.2</td>
<td>21.4</td>
<td>1.5</td>
<td>28.3</td>
<td>3.0</td>
<td>8.0</td>
<td>13.5</td>
<td>6.5</td>
<td>6.5</td>
<td>5.5</td>
<td>5.0</td>
<td>14.5</td>
<td>5.5</td>
<td>11.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>31</th>
<th>32</th>
<th>33</th>
<th>34</th>
<th>35</th>
<th>36</th>
<th>37</th>
<th>38</th>
<th>39</th>
<th>40</th>
<th>41</th>
<th>42</th>
<th>43</th>
<th>44</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>5.1</td>
<td>4.8</td>
<td>5.8</td>
<td>5.2</td>
<td>13.8</td>
<td>16.1</td>
<td>4.0</td>
<td>3.1</td>
<td>2.2</td>
<td>3.2</td>
<td>36.8</td>
<td>8.1</td>
<td>3.1</td>
<td>2.6</td>
<td>2.8</td>
</tr>
<tr>
<td>Valid</td>
<td>4.9</td>
<td>4.8</td>
<td>5.8</td>
<td>5.2</td>
<td>13.9</td>
<td>7.2</td>
<td>4.1</td>
<td>3.1</td>
<td>2.1</td>
<td>1.5</td>
<td>36.9</td>
<td>8.1</td>
<td>3.0</td>
<td>2.6</td>
<td>2.7</td>
</tr>
<tr>
<td>Partial C.</td>
<td>4.8</td>
<td>6.2</td>
<td>4.7</td>
<td>4.9</td>
<td>13.0</td>
<td>7.3</td>
<td>3.3</td>
<td>0.8</td>
<td>3.7</td>
<td>21.5</td>
<td>35.4</td>
<td>10.1</td>
<td>9.1</td>
<td>3.9</td>
<td>7.1</td>
</tr>
<tr>
<td>Incomplete</td>
<td>33.0</td>
<td>7.0</td>
<td>17.5</td>
<td>5.5</td>
<td>9.0</td>
<td>164.6</td>
<td>0.0</td>
<td>8.0</td>
<td>-2.0</td>
<td>0.0</td>
<td>21.0</td>
<td>6.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>47</td>
<td>48</td>
<td>49</td>
<td>50</td>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
<td>56</td>
<td>57</td>
<td>58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>2.7</td>
<td>9.6</td>
<td>3.4</td>
<td>4.3</td>
<td>2.6</td>
<td>2.4</td>
<td>3.2</td>
<td>3.0</td>
<td>3.9</td>
<td>7.6</td>
<td>3.2</td>
<td>2.5</td>
<td>3.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valid</td>
<td>2.8</td>
<td>9.7</td>
<td>3.4</td>
<td>4.3</td>
<td>2.5</td>
<td>2.5</td>
<td>3.1</td>
<td>3.0</td>
<td>3.9</td>
<td>7.5</td>
<td>3.1</td>
<td>2.4</td>
<td>3.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial C.</td>
<td>-1.0</td>
<td>5.4</td>
<td>1.8</td>
<td>3.0</td>
<td>6.2</td>
<td>-4.7</td>
<td>4.3</td>
<td>2.0</td>
<td>4.6</td>
<td>10.9</td>
<td>7.3</td>
<td>8.4</td>
<td>4.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incomplete</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 6. Time spent on each question group by presentation format
Figure 7. Accumulate time span group by presentation format
Figure 8. Time spending on answering each question by completeness

<table>
<thead>
<tr>
<th>Item</th>
<th>Time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td></td>
</tr>
<tr>
<td>Valid</td>
<td></td>
</tr>
<tr>
<td>Partial C.</td>
<td></td>
</tr>
<tr>
<td>Incomplete</td>
<td></td>
</tr>
</tbody>
</table>
One thing we were unable to explain is why about 7 people dropped out around question item number 36, which was neither the starting point of a new page in the block form nor a difficult question to answer. It actually began to ask some easy Internet usage questions from item 35. A client side script execution difficulty was a possibility. Failure to execute the client side JavaScript which triggered the data submission function in the hidden page can cause data to scatter or be lost or even not get collected at all. Lost or scattered data might occur when the client machine or connection to the Internet is too slow to sequentially send out responses and may have corrupted the browser’s ability to continue handling the JavaScript request. Some users who disable the JavaScript executing support in their browser will also experience this effect.

**Out of order answering**

From the time log we collected when a respondent answered each question, we found that some latter questions were answered before some items in the front. It caused the graphs we generated to have some “negative time.” We calculated the time spent on each question by subtracting the time log of the current question we were analyzing to the time stamp of the previous question, so the time spent on each question was not guaranteed to be positive. If one respondent decided to answer some question in the back then come back to answer a question in the front, “negative time” will occur. Here, we called this “out of order” answering. This finding was not part of our original experimental plan, but by adopting this new experimental method, this finding shows that we were able to examine some interesting survey answering decisions on the Internet survey. It is a valuable finding which improves the usefulness the automatic responding behavior observation program we designed for this study.
Figure 9. Drop out point versus total time spent
CHAPTER SIX
DISCUSSION

In this study, we found some valuable information. We can discuss these finding based on two aspects. The first aspect is based on the implementation of the automatic responding behavior observation program designed for this research. Another is from respondents’ responding behavior information we collected in testing the potential factors’ effectiveness. Before we start the discussion about the study topic, it is a good idea to reveal the findings about the similarities and differences between the traditional and the Internet survey from this research.

Traditional and Internet Survey

From this study, we found that the suggestions and experiences like survey content, sampling and research conducting procedures should still be applicable. Also, incentive is still a very important factor in influencing the rate if return in Internet surveys. The experience and knowledge from conducting surveys doesn’t fade away because this emerging new communication medium. In fact, previous research about the Internet in the communication field suggested that the Internet, as the newest communication channel in the world is actually no different from the life cycle of other channels (Lock, 1995; Morris, 1995; Newhagen, 1996). It started slowly and many people did not find out about it in its early stage. Then it exploded after it reached the critical mass. Now that almost everybody is online, it has managed to change people’s life style and researchers find that they still don’t fully understand how to properly utilize it. It tells us that even though people use the media differently now, the mechanism of how we can influence an audience with advertising and persuasion theories are still valuable knowledge for the
new platform.

On the other hand, we also found that some characteristics that belong particularly to this new medium is due to how the Internet interacts differently with its audiences. For example the styles of an Internet survey, listing all questions at once or using the paging functionality, and the way we can observe people’s behavior at the other end of the virtual world. In the past, limited to the constraints of all kinds of communication channels, practitioners and communication researchers likely suffered from the delay of slow communication tools (like mail, telegraph), lack of interactivity (like television, radio) and limited bandwidth of information transmission. But with the still improving Internet technology, people can now utilize all of the above functionalities in one place. Some popularly used methods like RDD and random sampling are actually difficult to implement in Internet research. Also, with the help of the Internet, the information transmission became extremely easy, but with the massive information available online, to promise that information sent will not be ignored is not an easy task. Hence, the survey invitation method proves to be an important factor in drawing respondents to Internet surveys.

**Response Rates and Drop Out Rates**

Response rates and non-response rates were the indexes generally used in previous Internet research to evaluate how well the survey was deployed. The response rate is defined in the traditional survey methods as the percentage of responses received from the total invitations sent. The non-response rate, in contrast, is defined as the percentage of respondents who didn’t return the survey. They not only indicate how well a survey is designed, but also determine the degree that the data can be trusted using
quantitative analysis. However, we argued that these definitions of the response rate and
the non-response rate are not enough for survey methodology analysis. It is highly
possible that between responses and non-responses, some respondents stop responding at
some steps in between before submitting the survey. The result of our experiment
demonstrated this.

Here, we referred to the non-responses as those who didn’t take any action at all
after they received the survey information. The dropouts, on the other hand, are those
who started the survey, but didn’t complete it. The response rate should include all
respondents who at least started the survey. If we say that the non-response rate provides
us with information about how effective the invitation method or message is, then the
drop out rate provides the clues to what survey presentation or structure makes
respondents decide not to finish the task.

Limited by the difficulty of collecting data before respondents click the final
submit button, the drop out rate is ignored in most of the research that uses the Internet as
a deploying platform. If we compare the drop out rate with the response rate or
non-response rate, the drop out rate that sits in between is hard to observe using
traditional methods. In this study, we questioned the use of the non-response rate as the
only measure to determine the quality of an online survey design and decided it was
inappropriate. We hypothesized that the drop out rate is also an important part of all of
the incomplete cases and have not been carefully analyzed in the past. We also suggest
that if non-response rates with dropouts and without dropouts are different, then looking
at drop out rates separately from non-response rates is necessary for future Internet
research.

However, our data did show different results on the hypothesis concerning the use
of different survey deploying strategies. When we compared the drop out rate with the non-response rate, the drop out rate was apparently more important than the non-response rate for groups that received higher response rates, such as the group that was invited in the middle of some other online task. But it was relatively not important if the survey itself (or the invitation method, reward provided) does not attract enough respondents to take the action to start it. Comparing the 33 incompletes (dropouts) with the 815 respondents who didn’t take any action, the number of incompletes was not good enough for us to use in support of our hypothesis. In this case, those who dropped out were relatively not important, and attracting more respondents to the survey becomes the better direction to proceed in.

On the other hand, the respondents who were rewarded with extra credit in class had a very high rate of responding to the survey upon perceiving the survey information. In this case, the drop out rate played a more important role when analyzing the survey deployment procedure for collecting more complete responses. In addition, when we inspected this by the invitation method, we found that the respondents who were invited while taking an online quiz, had a surprisingly low non-response rate (only 2%). The drop out rate therefore became the major factor that influenced the completion rate. In both cases, the problem of how to keep the respondents in the survey was the key to more valid responses. The analysis of whether to treat the drop out rate separately from the non-response rate provided us with a better understanding of how to accurately measure the problem of the procedure of deploying an online survey. More reward incentive or better invitation strategies should be implemented if the non-response rate is higher than the drop out rate, otherwise, improving the survey content to keep people in the survey should be investigated.
From the previous analysis, we concluded that attracting respondents and keeping respondents in the survey are both important, but the ratio of how important they are in an online survey depends on whether it was the reward or answering this survey that was appealing or if answering the survey was an easy task. Of course, we must first improve the attraction of the survey. If the survey topic or the reward can not attract more participants, the non-response rates will go up. So no matter how easy it is to answer the survey, the response rate will not improve. If, on the other hand, the survey itself can be attractive enough and more people respond to it, the method we use to keep them in the survey becomes more important.

**Factors that influence survey respond**

In this study, we tested three different factors that might influence the response rate. The first was the invitation method, the second was the reward system, and finally the survey presentation design. Based on this study, survey invitation methods and rewards appear to be the factors that will influence the motivations to respond to the survey (non-response rate), whereas the survey presentation styles influenced the respondents’ perception of how easy it was to complete the survey (drop out rate). However, our data showed some presentation techniques, like grouping questions into blocks and the use of a progress indicator can actually keep respondents in the survey longer for a higher completion rate. This suggests that perception time is not necessarily equal to the real time in responding to a survey. If we could find a better way to ease a respondent’s anxiety about finishing the survey, we would have a way to lower the drop out rates.

Previous research did not focus on survey invitation methods. They may believe
that distributing an invitation is nothing more than spreading out survey information. Serious researchers focus more on sampling to avoid the sampling error, but tend to ignore that a bad distribution design might cause a very high non-response error that eliminates the effort put into sampling. Some researchers tried to make their online survey reach as many audiences as possible by posting it on a popular heavily trafficked website. But more response doesn’t always mean you have representative data. A widely distributed survey without appropriate access control will introduce deviation by treating duplicated responses as two different responses. In addition, sampling error might cause the whole survey to be meaningless. This study shows that the survey distribution method plus an appealing reward can make some difference.

The finding that respondents who were invited to the survey while performing some other tasks online drew a very high response rate gave us some more information about online survey distribution. Because most previous Internet research only reaches a 30% response rate using regular e-mail invitation no matter the kind of reward, it appears that the reward itself is no longer a key issue for regular Internet users. Even if they are attracted by the reward, the e-mail invitation might be blocked by anti-spam software, they might forget the deadline, they might have more important things to do at the moment …etc. There are too many possibilities that the e-mail invitation, or the advertisement posted on the web page will be ignored in this world full of information. The invitation method, which invites the target while they are in the middle of some online task, might be a good strategy for surveying a targeted group. It would be even better if we could also include relevance to the task they are performing for participating in the online survey. For example, a 10% coupon offer for taking a customer satisfaction survey during their procedure of checking out after making an online purchase. The
reward could be used immediately and they would not miss this information because you are offering them a better deal on their current purchase.

In this research, we show that the invitation method and the reward method are two factors that can be manipulated to improve a sample’s intention to participate. A better survey presentation style, on the other hand, can keep respondents in the survey longer. By performing an analysis on the exception, the relatively small fraction of dropouts make the difference between each group that is too small to analyze using a formal statistical method. In addition, the result of our experiment show that with the non-response rates ranging from 1% (invited during the online quiz with the extra credit reward) to 87% (invited by e-mail with a free meal as a reward), the dropout rate appeared to be more stable across the groups (range 3.7% to 14%). The only factor we tested that should influence a respondent’s intention to drop out was not supported due to the low number of dropouts collected. This probably means that those who decide to respond will usually finish the task, unless the task is perceived to be far too difficult. However, even though it took longer for the respondents who received the survey partitioned into several pages with the use of a progress indicator to finish a survey than those who received the survey with all the questions on one page, the former survey style retains the lower dropout rate. This finding implies that we can improve the survey completion rate by adopting better and less burdensome survey designs. However, limited by the uneven nature of respondents in this study which uses mostly higher Internet users, we should continue to explore these potential factors with a more even user sample.

**Internet Responding Behavior Observation Program**

The most exciting finding in this study is the success of using the program we
designed to collect the data and observe the respondents’ responding behavior. The program was triggered by the “on-change” event of the data field of the survey page. It helps us to collect responses along with the responding time by sending small pieces of data to the server in the background. By collecting the exact time respondents responded to each question item, we could evaluate how long on average it took to answer each question. In addition, we found the dropout point by using the time a respondent responded to the last question item if the submit button was not clicked (no survey end time was recorded). Some detailed information like what distracted them from answering the survey or what content in the survey discouraged them from finishing the survey is still not collectable using this program. Also, the last time stamp we collected was not necessarily the last minute they spent on the survey. However, it is these characteristics of Internet surveys that allow Internet users to browse around without any limitations. Because each movement during responding was recorded, the data this program collected should be enough to analyze frequent drop out points or other survey responding related factors such as out of order answering.

From our analysis, certain common dropout points were found in our question set which is a common survey size. We found that the point that most respondents dropped out is when we started asking scale questions that required some thought. Of course, for the respondents who dropped out, the more questions they answered the longer they spent answering. It is reasonable that as respondents approach the end of the survey, they most likely won’t drop out of the survey. Our result also shows that all dropouts happened before question item 42. Hence, the question becomes: should more difficult (or sensitive) questions be located at the end of a survey? Comparing the Internet survey with the traditional mail survey, the effort to submit an Internet survey is less than to submit a
Will it be easier to collect responses only on certain types of questions on the Internet? Further research is needed to find the common drop out point patterns and a better question ordering strategy by testing the questionnaire with a different order of questions.

By analyzing the time spent on each question, two types of spikes in the time-question figure are worth an extra look. The first is the spike that happened across the survey presentation format. It means that respondents need more time to respond to that specific question no matter what style we used to present the questionnaire. Another type is the spike that appeared only in a certain presentation style. The appearance of the former type of spike indicates that the longer response time is caused by some systematic issues, and the appearance of the latter type of spike shows that some style specific issues might need some further inspection. The finding again proves the value of using the observation program. By knowing this information, analysis on a survey could be performed in a more detailed way, because we would not be confused by whether the problem is in the question or the presentation style. Hence, the problem can be easily identified.

This research is the first step in collecting information about a respondent’s decision making process while taking an online survey. Even thought we developed the technique to collect step by step information in this experiment, our results show no extra cost to use this while collecting survey responses. Hence, this method is safe to run parallel with online surveys. Furthermore, by using the data collecting program, we found that the extra information collected is extremely helpful for survey researchers in understanding the quality of their online survey design. For example, the dropout points
and time spent on each question provide Internet survey conductors a great deal of information about the problem with their survey, and the out of order answering also reveals the problem of particular questions or the order of them. Hence, no matter how much Internet survey conductors care about the completeness or the drop out rates of their Internet surveys, this method is highly recommended to use for pretesting their online surveys before deploying. If a well designed pretest is performed, the weaknesses in the questionnaire will be recognized before being formally conducted in the real world.

Besides its use as a pretest tool of Internet surveys, it can also help real Internet surveys collect partial data even if the respondents did not submit. One thing that should be noted is the ethical issue of collecting partially completed responses. Could we use this partially collected data with the program we designed? If future studies want to be able to utilize this functionality, a clear statement about when they will begin collecting responses should be included in the invitation message or consent form. In addition, the program just provides an inking or insight into the capability of the Internet. Because of the fast pace of the developing Internet technology, the same functionality and JavaScript code is not always guaranteed as a future standard. Modification and adjustment might be required for future implementation.

**Conclusion**

From this study, we found that even though the presentation style is not an issue in traditional survey methods, the use of a plain HTML page for survey data collecting is actually not a very good idea. Like what we revealed, even though groups in this experiment didn’t show serious differences in response rates, you can not ignore the dropouts in Internet surveys, and this kind of ignorance should still not be tolerated.
Continuing research is suggested in this study in order to build more understanding of the mechanism of a respondent’s Internet survey responding process, and the knowledge about how to hold audiences longer in Internet surveys, or any virtual place.

Apparently, the drop out rate we drew was not high enough compared to the non-response rate, but one thing worth considering is that this was a well designed questionnaire that contained only 58 question items and has been tested in the traditional survey platform. It is a reasonable assumption that more question items would lead to more dropouts. What we can do to inspect this assumption is utilize the program with a collection from a question bank, and present a different number of questions randomly to every participant to test the tolerance of survey length. By quantifying survey length as an independent variable, we would have a better understanding of an appropriate Internet survey length. Also, strategies to deploy online surveys can have more effect in gathering more responses and keeping respondents throughout the survey.

Though the burden of an Internet survey is reflected by the drop-out rate pulled using our program, this research methodology is relatively new to Internet research, and this study was performed on a relatively homogeneous population, so we can not be assured that our finding is strong enough to show a factor’s influence on a respondent’s Internet survey responding process. Further testing on a broader and more heterogeneous population, especially on both heavy and light Internet users of all educational backgrounds, with this program is recommended. However, we did show that incentive is important to Internet surveys as it is important to traditional ones, and the best result can be achieved by providing a meaningful incentive to the right target with the right method.

Finally, the more valuable finding with the exception of the test of the potential factors in the survey responding process, is the program we used to collect this extra data.
Our study indicates that the program is suitable for all Internet survey pretesting. Actually, with the ease of this program for respondents with appropriate server side design, the interactive Internet survey or Internet communication channel can be implemented. Here, we have successfully showed that we can delivered the idea of observe responding behavior online. With the support of real time response and automatic branching, it can even be modeled as a vivid Internet interviewer that brings us to the next generation survey platform.
BIBIOGRAPHY


Bachmann, D., Elfrink, J., & Vazzana, G. (1996). Tracking the progress of e-mail versus snail-mail. Marketing Research, 8, 31-35.


Schneider, S. J., Kerwin, J., Frechtling, J., & Vivari, B. A. (2002). **Characteristics of the discussion in online and face-to-face focus groups.** *Social Science Computer Review*, 20(1), 31-42.


Smith, C. B. (1997). **Casting the net: Surveying and Internet population.** *Journal of*


APPENDIX
Appendix A. Consent form

My name is Ching-Guo Wu, and I am a graduate student in the Department of Communication at Washington State University. I am conducting this research for my thesis regarding Internet usage, and I'd like to ask you for your help by answering an online survey for me. This study has been approved by the Institutional Review Board at Washington State University.

Your participation is highly appreciated. This survey will only take a few minutes to complete. Please complete this survey and click the finish button. Once you agree with the term and enter the survey by clicking the "Yes" button below. Of course, your participation in this survey is voluntary and your e-mail address will be removed to maintain anonymity. If you feel uncomfortable about this survey, you can stop answering questions at any time without penalty.

If you have any questions, feel free to contact me at cwu@wsu.edu. If you want a copy of this study results, just inform me through e-mail and I will send it to you.

Thank you very much for your participation and quick response.

Cordially,

Ching-Guo Wu
Graduate student in Washington State University

Are you ready to start the survey?
Appendix B. Survey questions

(R1) 1. What is your gender?
- MALE
- FEMALE

(R2) 2. What is your race or ethnicity?
- AFRICAN AMERICAN/BLACK
- NATIVE AMERICAN
- ASIAN AMERICAN
- EURO-AMERICAN/WHITE
- HISPANIC
- OTHER:

(TB1)

(R3) 3. What is your age?
(Survey questions continued)

4. How many minutes did you spend reading any newspapers yesterday?

5. How many hours did you spend watching TV yesterday?

6. About how many friends or family members did you talk with yesterday?

7. About how many other people did you talk with yesterday?

8. How many times a week do you watch a local television newscast?

9. How many times a week do you watch a national television newscast?
(Survey questions continued)

10. How accurately would you say the local television news teams present the events they report, would you say very accurately, somewhat accurately, somewhat inaccurately or very inaccurately?

- VERY ACCURATELY
- SOMEWHAT ACCURATELY
- SOMEWHAT INACCURATELY
- VERY INACCURATELY

11. How accurately would you say the national television news teams present the events they report, would you say very accurately, somewhat accurately, somewhat inaccurately or very inaccurately?

- VERY ACCURATELY
- SOMEWHAT ACCURATELY
- SOMEWHAT INACCURATELY
- VERY INACCURATELY

12. How often do you use e-mail? Would you say almost everyday, weekly or monthly?

- EVERY DAY
- WEEKLY
- MONTHLY

(Survey questions continued)
<p>| (R13) It's intimidating to learn Internet use. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R14) Internet literacy makes work easier. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R15) The cost of Internet software is too high. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R16) I am fearful of using the Internet. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R17) Internet literacy offers new opportunities. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R18) I think technology destroys our society. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R19) I am willing to learn new ideas. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R20) I believe the Internet is a useful information tool. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R21) I am willing to take a risk. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R22) It's frustrating to learn Internet use. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R23) I am excited to learn more about the Internet. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R24) I think technology such as the Internet makes people antisocial. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R25) It's difficult to learn Internet use. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R26) There is no privacy on the Internet. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |
| (R27) Internet literacy is needed at work. | ![ ] | ![ ] | ![ ] | ![ ] | ![ ] |</p>
<table>
<thead>
<tr>
<th>Row 1</th>
<th>Row 2</th>
<th>Row 3</th>
<th>Row 4</th>
<th>Row 5</th>
<th>Row 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>The cost of Internet access is too high.</td>
<td>I like to keep up with new technologies.</td>
<td>I think technology such as the Internet is good for our society.</td>
<td>Internet literacy makes life easier.</td>
<td>The cost of a personal computer is too high.</td>
<td>I am willing to explore new technologies.</td>
</tr>
<tr>
<td>I believe there is perverted content on the Internet.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Page 4
(Survey questions continued)

14. Would you consider yourself a heavy, medium or light Internet user?

- HEAVY
- MEDIUM
- LIGHT

15. How many days in the last week have you used the Internet?

16. Where do you have access to the Internet?

- WORK
- SCHOOL
- HOME
- OTHER:

This page continues the survey questions.
(Survey questions continued)

17. Where do you use the Internet?

18. How knowledgeable about the Internet do you consider yourself? Would you say extremely knowledgeable, very knowledgeable, somewhat knowledgeable or not knowledgeable at all?

EXTR EMELY

VERY

SOMEWHAT

NOT AT ALL
(Survey questions continued)

<table>
<thead>
<tr>
<th>Question</th>
<th>Option 1</th>
<th>Option 2</th>
<th>Option 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>19. Do you own any of the following?</td>
<td>YES</td>
<td>NO/DON'T</td>
<td>KNOW</td>
</tr>
<tr>
<td>(R42) Personal computer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R43) Satellite dish</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R44) VCR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R45) Digital video disk or DVD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R46) Video camera</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R47) Compact disk player (CD player)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R48) Laser disk player (LD player)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R49) Video game console</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R50) Electronic personal organizer (PDA)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R51) Electronic pager</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R52) Answering machine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R53) Cellular telephone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R54) Fax machine</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Page 7
(Survey questions continued)

20. Do you subscribe to the following?

<table>
<thead>
<tr>
<th>Service</th>
<th>Yes</th>
<th>No/DON'T KNOW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic cable TV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premium or pay cable TV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DBS or DDS satellite service</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voice mail</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Page 8**