This is to certify that the
thesis entitled

ONLINE MOOD ASSESSMENT: IS DAILY MONITORING A REACTIVE MEASURE?

presented by

Paul Edward Quinlan, DO

has been accepted towards fulfillment
of the requirements for the

Master of Science	degree in	Epidemiology

Major Professor's Signature

7 May 2007

Date

ONLINE MOOD ASSESSMENT: IS DAILY MONITORING A REACTIVE MEASURE?

By

Paul Edward Quinlan, DO

A Thesis

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Master of Science

Department of Epidemiology

2007

ABSTRACT

ONLINE MOOD ASSESSMENT: IS DAILY MONITORING A REACTIVE MEASURE?

By

Paul Edward Quinlan, DO

This R01 application combines the strengths of faculty at Michigan State University (MSU), Johns Hopkins University (JHU) and the University of Vermont (UVT) to evaluate if online daily monitoring of mood, by itself, might be a reactive measure. College-age young people recruited as anonymous study participants will engage in daily or monthly online mood monitoring with the MSU online 'Longitudinal Study Engine' (LSE). The LSE automates initial evaluation and longitudinal follow-up of a participant's response to mood assessment permitting immediate feedback to the participant. Because some young people are expected to deviate from the prescribed online daily or monthly monitoring, inference from study evidence about a suspected effect of daily monitoring is complex. Joint collaboration with JHU and UVT will promote the project's success in spite deviation in mood monitoring. The project promises important new evidence about online monitoring of mood in young people.

DEDICATION

I dedicate this thesis to my wife, daughters, family, relatives, friends and colleagues.

ACKNOWLEDGEMENTS

Firstly, I would like to acknowledge the support provided by my thesis guidance committee members at Michigan State University in the College of Human Medicine. The thesis guidance committee members from the Department of Epidemiology, Professor Naomi Breslau and Chair James Anthony have been instrumental in my effort to complete this research application as my thesis. Their intellectual support and academic guidance has been invaluable to my work. The thesis guidance committee member from the Department of Psychiatry, Chair Jed Magen, has provided intellectual and academic support as well as accommodation of my time for completion of coursework for a Master of Science in Epidemiology. Collectively, the varied backgrounds of my committee members provided insight into the development and realization of this application.

Secondly, I would like to acknowledge the role of my instructors in the Department of Epidemiology. Their academic efforts would not have made this grant application possible.

Lastly, I would like to acknowledge the support of additional faculty at MSU and elsewhere. Dr. Carlos Rios-Bedoya provided guidance to me in the grant application preparation process and assisted me in grant preparation at a time when I was uninitiated in the process. All of the faculty and students of Dr. Anthony's research team working on the implementation and application of the Longitudinal Study Engine have been instrumental to this application's

completion. I wish to specifically acknowledge and thank Dr. Jodi Holtrop of the Dr. Anthony's LSE research group. The design and development of Drs. Holtrop and Anthony's online reactive measure study in smoking cessation with the LSE was critical to the development of this application. I also wish to acknowledge Dr. Constantine Frangakis at JHU and Dr. John Helzer at UVT for their agreeing to provide their support for this application. I would like to acknowledge the clinical faculty at MSU for supporting this grant application through insuring the safety of future participants: Dr. Glynda Moorer and Dr. Leigh White at the MSU Olin Student Health Service and Counseling Center and Dr. Christopher Giuliano in the Department of Psychiatry at MSU.

PREFACE

The Department of Epidemiology at Michigan State University allows the preparation of a research proposal as a method for completion of the thesis requirement for the Plan A Master of Science in Epidemiology. For my thesis, I have chosen this method.

The result of the development of this thesis is that it has been submitted as a R01 grant application to NIH in a response to Program Announcement PA-07-070. This application was submitted to NIH in March 2007.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

$	U.S. Dollar
AUDIT	Alcohol Use Disorders Identification Test
CACE	Complier Average Causal Effect
CES-D	Center for Epidemiological Studies – Depression scale
Co-PI	Co-Principal Investigator
CV	Curriculum vitae
DSM	Diagnostic and Statistical Manual
DSMB	Data Safety Monitoring Board
HIPPA	Health Information Portability and Privacy Act of 2001
HIV	Human Immunodeficiency Virus
I	Index
ID	User Identification
IRB	Institutional Review Board
ITT	Intention to Treat
IV	Intervention
JHU	Johns Hopkins University
LCM	Life Chart Method
LSE	Longitudinal Study Engine
MSU	Michigan State University
MTF	Monitoring the Future Survey

N	**Sample size**
NEP	**Needle Exchange Program**
NIAAA	**National Institute on Alcohol Abuse and Alcoholism**
NIDA	**National Institute on Drug Abuse**
NIH	**National Institute of Health**
NIMH	**National Institute of Mental Health**
PI	**Principal Investigator**
PS	**Principal Stratification**
PTSD	**Post-Traumatic Stress Disorder**
RA	**Research Assistant**
TM	**Trademark**
UCRIHS	**Michigan State University Committee for Research Involving Human Subjects**
URL	**Uniform Resource Locator**
UVT	**University of Vermont**
WHO	**World Health Organization**
WWW	**World Wide Web**
YST	**Youth Nominated Support Team**

CHAPTER 1: SPECIFIC AIMS

In this R01 research project, we are attempting to build a new methodology for experimental research on health interventions that can be implemented on a mass scale. The proposal is responsive not only in its presentation of an innovative methodology for gathering fine-grained anonymous longitudinal data about mood status (e.g., day-by-day reporting of mood via an online Internet log-in session by anonymous respondents), but also in its objective of tailoring the principal stratification framework for causal inference to the specific inferential challenges that are faced in longitudinal and intervention research of this type (e.g., non-compliance with prescribed experimental regimens; non-ignorable missing data).

Our MSU research group has developed the Longitudinal Study Engine (LSE) as an online tool for completion of anonymous longitudinal studies. Within LSE we can embed randomly assigned online interventions (e.g., flash videos, interactive health-oriented games) as well as randomly assigned incentives to begin and to sustain experimental control over subject participation levels. Our intent is to make the LSE available for NIH research on a not-for-profit basis, as has been done for NIH proposals submitted by faculty members at Michigan State University (MSU) and elsewhere (e.g., Dr. Jennifer Havens of the University of Kentucky).

Nonetheless, before using the innovative LSE approach to evaluate more

intensive online interventions (e.g., flash videos) we must address an important research issue recently raised in the NIAAA-sponsored research of John Helzer's University of Vermont (UVT) research group. Namely, when we encourage daily monitoring versus monthly monitoring, is daily monitoring via the Internet a reactive measurement with respect to target outcomes such as improvement in mood? Work of Helzer's UVT research group suggests that daily monitoring, by itself, might have reactivity with respect to drug-taking behavior (Helzer 2002). We have engaged John Helzer, a long-time collaborator, to join us in this research project so that he might share his experience and expertise in a deliberative way.

In the future, our intent is to use online daily monitoring of mood as a tool to evaluate the impact of more intensive interventions, with LSE-enabled daily monitoring and delivery of the online interventions. However, before that step in the research program, we must make an experimental contrast of what happens when we encourage experimental trial participants to engage in daily versus monthly monitoring. To this end, we are proposing a monitoring trial, with randomization of participants to conditions of daily versus monthly monitoring. Before we extend the sampling frame to more general population segments, we will conduct the trial with a population sample of college-attending young people.

In this step of the research program, we face some interesting and thorny inferential challenges in relation to the concept of 'potential outcomes' that was introduced by statistician Jerzy Neyman more than 50 years ago. Namely, when we encourage daily monitoring in one randomly assigned arm of the study, some

of the participants will not comply. Some of the potential participants will not log-in at all, despite incentives and encouragement to do so. Others will log-in to initiate daily monitoring, but will not demonstrate 100% compliance with the daily monitoring prescription. For example, they might use a wall calendar to keep track, day by day, of their mood in their own jury-rigged version of our online daily monitoring.

This type of non-compliance introduces complexity in the inference of the causal effects of the randomized conditions. There is a rich tradition of biostatistical methodology research on this topic, recently summarized in a series of influential papers authored by Constantine Frangakis and Donald Rubin on the topic of the 'principal stratification' framework for causal inference (e.g., Frangakis & Rubin, 1999, 2002). In order to address these complexities, our MSU research team already has engaged Dr. Frangakis as a collaborator and co-investigator in related research, under a subcontract financed with MSU funds. This research project and collaboration is described in Chapter 3 under 'Preliminary Studies.' We propose continuation of this collaboration with Dr. Frangakis as part of the proposed R01 research project, in order to ensure his involvement in the tailoring of the 'principal stratification' framework to the specific contours of this type of randomized trial design, with participants randomly assigned to the daily monitoring arm contrasted with participants assigned to the monthly monitoring arm. As such, the tailoring of 'principal stratification' represents a methodological advance and innovation that complements and strengthens the advances and innovations represented by the

LSE methodology for anonymous longitudinal assessment and delivery of online interventions.

Primary Aim

1. The primary specific aim of this research project is to answer this research question: "When we encourage daily monitoring versus monthly monitoring under conditions of a randomized trial with tangible incentives for participation, is daily monitoring via the Internet a reactive measurement with respect to target outcomes in mood?"

Hypothesis for Primary Aim

1. The online monitoring of mood is reactive in its effect on the study participant's mood state with daily monitoring having a greater effect on mood than monthly monitoring.

Secondary Aims

1. In an effort to plan for future randomized trials and observational studies with online monitoring, a secondary aim is to estimate survey response levels that can be secured with pre-specified tangible incentives for participation.
2. In an effort to extend the methodological innovations represented by this project, another secondary aim is to tailor the principal stratification framework for causal inference to the context of observational

epidemiological research with daily and monthly monitoring of mood as well as the context of a randomized trial to evaluate alternative and potentially reactive online interventions.

CHAPTER 2: BACKGROUND AND SIGNIFICANCE

SECTION 2A: Background

This research is motivated by an effort to foster NIH-supported methodologies for developing and evaluating health behavior and lifestyle change modalities that can be delivered on a mass action scale — with 1000s or even 10s of 1000s of beneficiaries — conducting epidemiological surveillance on the target outcomes of interest on this large sample scale. In the process, the methodologies must be devised and refined so as to produce evidence that will support rigorous causal inference about the intended beneficial effects (and unanticipated harmful effects) of the modalities under evaluation. Here, we are proposing research on a methodology that harnesses the power of the Internet in order to gain affordable increases in the numbers of research participants in the evaluation studies, with the interventions delivered via online methods so that the numbers of potential beneficiaries can be increased. The target outcomes of this initial project involve assessment and reporting of a participant's mood state. However, as this program of research develops, our expectation is that other facets of mood disorders and related symptoms such as suicide will be substituted for this project's mood monitoring focus.

With respect to background and significance, depression represents a global public health problem, with millions affected worldwide. Depression afflicts individuals across the lifetime. Within the college age population, academic

performance worsens for those with mild depression. In more severe presentations of depression, college students are at risk of suicide (The American College Health Association). Depressive symptoms are present in nearly all college students who attempt suicide (Kisch 2005). Individuals with depressive symptoms who seek help have relied on self-initiated or externally encouraged contact with health and mental health professionals. One third of the U.S. population age 18-24 is either part time or full time college students (http://www.census.gov/prod/2005pubs/p20-554.pdf). This large cohort has Internet access and would be an ideal population to assess if online monitoring of mood is feasible and reactive in changing mood.

In addition to the above, the background of this project involves five topics. First is the magnitude of health problems due to depressed mood, as described in Section 2.B. Second is a potentiality for mood change that might be induced via online interventions described in Section 2.E. Third is the possibility that epidemiological surveillance via the Internet is feasible and that daily monitoring of mood might induce a tangible impact upon the health behaviors of interest as described in Section 2.I. The fourth topic involves the biostatistical approaches required to draw causal inferences from observed study data of the type proposed for this research project as described in Section 2.J. The fifth topic involves a rationale for starting with college students and other young people in a program of evaluative research into online interventions as described in Section 2.K. Section 2.L provides a brief overview of the public health significance of the proposed R01 research project.

In these sections, we seek to build a case for NIH support to complete probing research into the possibility that daily monitoring of mood might, by itself, constitute a reactive measurement, with its own tangible effects on the level of the participant's mood state. The thrust of our argument is that fine-grained data on the mood state is crucial if we are to understand the mechanisms through which online intervention programs are influencing levels of mood, where 'fine-grained' data on day-to-day experiences may be contrasted with 'coarse-grained' data gathered retrospectively on month-to-month or year-to-year experiences. Nonetheless, in the process of gathering the fine-grained data, we may be influencing the mood state of the participant that we seek to influence via online interventions of greater intensity (e.g., flash videos, real-time online cognitive behavioral treatment of depression, or the other forms of online intervention described below).

SECTION 2.B: Lifestyle As a Major Determinant of Health

For centuries, an individual's change in lifestyle in response to mood state has been recognized as major contributor to ill-health. The 20th century provided scientific evidence that indicts specific facets of lifestyle as causal determinants of ill-health and premature mortality — especially with respect to mood state, but also with respect to sexual practices, diet, and exercise.

The potential public health significance and scientific merit of the proposed R01 exploratory and developmental research project rest in large part upon its potential to harness the power of online epidemiological surveillance and online

interventions in the service of lifestyle changes that can lead to lasting health benefits and resulting extended survivorship. In this context, we have considerable optimism that online stimuli and programs can be automated on a mass scale for relatively low-cost interventions intended to reduce the occurrence of hazardous changes in lifestyle as can occur in a depressed mood state — i.e., at a cost that is a small fraction of the costs of interventions requiring person-hours of skilled professional time. This optimism is bolstered in part by a tremendous acceleration in the number of public access points for online interventions — an acceleration seen even among the most resource-challenged of the world's population. Within many emerging market economies, for example, cybercafé access points have proliferated and the cost of online sessions has dropped to affordable levels. For example, within Mexico and Peru, there are cybercafés or Internet access points for large proportions of the population in smaller villages as well as major metropolitan areas; supply of these access points has resulted in very favorable costs per log-in for individuals who wish to tap the power of the Internet. Within the United States, well over 50% of households have regular Internet access; the expectation is greater than 90% within 5 years.

The importance of lifestyle change for health improvement is conveyed in the World Health Organization's goals for the future of its public health work, and in the health goals specified by the U.S. government for this population. The number of beneficiaries who qualify for this coverage already far outstrip the

funds allocated to finance coverage of standard telephone help-line counseling programs.

Due to the reduced cost advantages of automated online interventions for depression, it is important to develop and refine online methodologies for large-sample epidemiological surveillance of mood states, as well as complementary methodologies for delivery and evaluation of online interventions. The question is not whether online interventions will become part of the future of public health work in the United States and other countries. Questions for the future of public health research and public health response concern when and how effectively these online interventions will be initiated to promote change in mood states.

The proposed research is a step in a more general program of health promotion research that is intended to harness the power of online tools for epidemiological surveillance and for evaluative research to compare and contrast alternative online tools to promote improvement in mood. Depressed mood represents an early target in our research program on this front. The results of the proposed research on mood will help us to guide the more general research program on health promotion.

SECTION 2.C: Potentiality for Lifestyle Change

Public health researchers already have started to harness the Internet to deliver online interventions directed toward improving mood, including changes in symptoms of depression. In many instances, these online interventions are building upon experiences with the evaluations of programs developed during

pre-Internet days. Examples of cognitive behavioral therapy programs are illustrative.

SECTION 2.D: Successful Mood Interventions

Before involvement of Internet methodologies, considerable progress was made in relation to interventions directed toward treating depression. Effective methods are available to assist individuals with depressed mood such as group or individual behavioral counseling. Among adults, Cognitive Behavioral Therapy, is effective in treating depression. Although issues have arisen in meta-analyses of the efficacy of antidepressant medications, serotonin specific reuptake inhibitors have been efficacious in treating depression in adolescents and young adults (March 2004, Lieberman 2005).

SECTION 2.E: Successful Online Interventions

The cognitive behavioral model for treatment of depression has been extended into the domain of online interventions. The MoodGym developed by the Australian National University (http://moodGYM.anu.edu.au) is an online cognitive behavioral treatment program developed to treat mood symptoms. Its interactive interface can assess online participants' mood symptoms and provide strategies to treat the symptoms.

SECTION 2.F: Longitudinal Study Engine (LSE)

Recognizing the potential for online epidemiological surveillance of large population samples and for the capacity to undertake automated randomization of online health promotion stimuli and programs, our MSU research group designed a Longitudinal Study Engine (LSE). The LSE has the capacity for (a) anonymous longitudinal log-ins with automated tracking of an individual's responses to standardized online survey assessments, made possible by allowing the sampled participant to draw a ticket number at random from among 1000s of ticket numbers, and then by presenting the participant with the opportunity to create an individualized anonymous user ID and password at the time of the online ticket redemption, (b) repeated log-ins of the individual, who uses the same self-generated anonymous user ID and password to re-enter the survey site and to complete the online assessments serially, (c) computerized adaptive mechanisms to allow individual-level tailoring of survey assessments as well as delivery of online interventions on a deterministic or probabilistic basis, and (d) use of the principles of behavior analysis to deliver incentives according to schedules of reinforcement that optimize the log-in and reporting behavior (e.g., via variable interval reinforcement schedules).

The LSE is the epidemiological surveillance platform that will be used in the proposed R01 exploratory and developmental research project and in our research program on the topic of online health promotion research. In brief, an epidemiological survey participant is recruited from a sampling frame; draws a ticket number at random from the 1000s of available ticket numbers; completes a brief pre-login survey that is used to record the randomly drawn ticket number

and to gather data about those who do and do not participate; logs onto the URL provided with the ticket number and keys in the ticket number, after which there is a prompt for self-generation of the unique userid and password, plus password memory prompts. At that point, the respondent is presented with an online disclosure statement for participation in the longitudinal and experimental elements of the research protocol; respondents who decline to participate are credited a gratuity value that can be redeemed for simply having logged into the LSE system; respondents who agreed to participate are credited the assigned gratuity value, plus the additional credit for a gratuity to thank them for completing the initial online survey. Thereafter, respondents are alerted to the possibility that they will receive additional gratuities for repeat log-ins according to the protocol schedule, and that they might be eligible for sub-studies under certain conditions. The LSE delivers the standardized survey assessments according to an investigator-specified algorithm, which can include algorithmically implemented branching patterns (e.g., computerized adaptive testing based on early test scores), as well as algorithmically delivered flash videos, video games, or other online stimuli intended to influence future participation and/or health promoting behavior.

As such, the LSE platform has been designed to facilitate and to automate large-sample epidemiological research (as well as small sample clinical studies), and to facilitate randomized trials of online interventions. The power of the LSE is being harnessed in this research project for work on a question of reactivity of daily monitoring of mood, in the service of a next step in the research program, in

which LSE-enabled daily monitoring of mood will be joined with an intensive Cognitive Behavior Therapy program in one arm of the study versus LSE-enabled daily monitoring alone in the another arm of the study. That is, once the issue of reactivity is settled, and reactivity has been estimated rigorously, we shall turn to the task of estimating the effect of online interventions in which the LSE-enabled daily monitoring of mood is combined with randomly assigned intervention and control conditions.

SECTION 2.G: Overview of monitoring of and reactive measures

The emergence of the Internet and its potential for health monitoring and intervention is of singular interest in this application. Careful consideration of its application in this role is necessary due to previously known issues in other methods of health monitoring. An individual's reactivity to measures has been understood for greater than half a century with traditional methods of collecting data. Reactivity to measure via health monitoring through the Internet cannot be dismissed without determining if can occur. In the next two sections, Internet health monitoring will be discussed followed by discrete examples of reactivity occurring in health monitoring via the Internet.

SECTION 2.H: Internet monitoring of health-related behavior

Internet and online survey methods offer a tremendous potential aid to the collection of lifestyle and health-risk behavior data during epidemiological surveillance and evaluation research. Internet monitoring already has been demonstrated to be feasible (Daley 2003; Pealer 2001; Pealer & Weiler, 2003,

Wong 2006, Santor 2007) and to carry many advantages. For example, surveillance by these methods often requires less time and is less expensive than telephone or mailed surveys. They also offer response option features such as check boxes, radio buttons, and text-entry boxes that pop up only in response to certain inputs/answers, with allowances for visual reminders and enhancements (e.g., the use of dynamically assigned image anchors for response scales as opposed to static verbal anchors). In some applications, email portals are used to recruit participants, and in this instance, disadvantages can surface, including the possibility that spam screening systems automatically delete unrecognized email intended to prompt responding (Pealer 2001).

Pealer (2001) compared web versus mail health behavior surveys of college students in Florida and found that students responded at similar levels to the Internet survey, but responded an average of 2.42 days sooner and skipped fewer sensitive questions on average than mail respondents. These results led the authors to believe that students are more likely to answer more sensitive items using online survey methods than a mail survey, and that response rates are similar. There were no more errors in the Internet than in the mail group. Overall, comments from the web group were overwhelmingly positive about the survey and resulted in a sixty percent response rate with an incentive of just $2.

With respect to response levels for online surveys, it is clear that response levels can be manipulated by the investigator. Several researchers have found that monetary incentives increase the response levels, but not necessarily data quality — regardless of the incentive delivery method (Birnholtz 2004; Bosnjak &

Tuten, 2003; Scholder 2001; Shaw, 2001). Within this line of research, the type and timing of the incentive has also been studied in relation to the response rate.

For instance, Birnholtz (2004) reported that cash incentives lead to higher response rates than gift certificates, either by mail or email. Among gift certificates, no differences in response rates were found between paper and electronic gift certificates. However, the sending of reminders was found to increase response levels, whether the reminders are sent through email or via paper invitations.

Another study looked at the effect of cash incentives on the response rate using three different delivery methods: by phone, by mail and by email (Scholder 2001). Mail and email surveys were offered a $0, $5, and $10 incentive to participate. At $5 incentive level, the response rate tripled for mail and email participants; the use of $10 incentives had greater incremental effect for mail than online participants (14 points vs. 7 points increase). Based on a study using randomly assigned cash incentives of $2 and $5 in a community mail survey, Shaw (2001) reported an increased response rate using $5 incentive with one survey mailing, when compared to $2 incentive with multiple mailings.

An additional issue has to do with the effect of prepaid versus postpaid incentives vs. prize draws. Prize draws used in web surveys have been found to be associated with increased completion and response levels when compared to postpaid ('promised') incentives (Bosnjak & Tuten, 2003; Downes-Le Guin, 2002; Tuten, Galesic, & Bosnjak, 2004;). Church (1993) found that cash and gift certificate incentives were effective only when included with the survey, with an

average increase of response rate of 19% for cash incentive vs. 8% for gift certificate. The diminishing return model proposed by Armstrong (1975), states that an asymptote will be reached: increasing the amount of the incentive should have a decreasing effect on response level once the asymptotical value is reached. In some applications, an increased incentive might increases response level in the first survey, but the effect might disappear for the following surveys (Kephart & Bressler, 1958).

Other forms of health monitoring, i.e. mail-in surveys, and direct interviewing have limits and benefits as survey methods. Online assessment by web survey has been studied by different scientific disciplines since the 1990s. The benefits of web surveys in ease of use and access to participants particularly in the wired and wireless generations offer great potential in an expanding era of information technology. One potential limitation is the perception by web survey participants that surveys should include a randomized incentive such as a lottery or prize drawing (Bosnjak & Tuten, 2003). Other earlier survey methods have not relied on randomized incentives as a method to recruit participation. Heerwegh found that within a college student population, prize drawings influence participation in web surveys (Heerwegh 2006). Heerwegh's analysis also suggested that some subgroups (males, positively influenced to participate, females no difference) within the survey are more influenced by the lottery than other participants. This pattern of participation due to a potential variability of the incentive (lottery, prize drawing, or random incentive) improves rates of response to web surveys. The variability of influence between subgroups needs to be

measured increasing the complexity of the analysis of a web survey. The limitation of increased complexity as perceived within the literature is a minor one compared to the achievement of greater participation by respondents.

With respect to anonymity versus confidentiality, the contemporary best advice is that a condition of anonymity yields more complete and accurate reporting than does a confidential condition – at least with respect to socially stigmatized and illegal behaviors (e.g., see Anthony, 2000). For this reason, the proposed research involves a completely anonymous condition; it is not possible for the investigators to know the identity of the individual respondents.

SECTION 2.I: Might daily monitoring be reactive?

Campbell defined this potential influence on internal validity as "A reactive measure is one which modifies the phenomenon under study, which changes the very thing that one is trying to measure." (Campbell 1957). The implementation of a methodology utilizing daily online access to record participant responses needs to be rigorously assessed. The rapid growth of online surveys in other disciplines heralds the potential application in mental health-related studies. It is unknown if daily online measures differ in their impact on internal validity as a reactive measure compared to other methods for survey, (e.g., mail, telephone, direct interviewing). These traditional methods of measure are not easily applied to daily monitoring. The feasibility for daily monitoring can be achieved online. We need to determine, however, if its influence can affect the validity of a study.

In our planning for the use of the LSE to evaluate online cognitive behavioral therapy for depressed mood, we had anticipated use of the LSE to secure a daily trace of the fine-grained health-related behavior of interest (e.g., daily mood state). Our discussions with NIH-funded collaborator John Helzer and his UVT research group increased our awareness that the daily LSE-enabled monitoring might in and of itself constitute a reactive measurement.

In summary, Helzer led a team of researchers at the University of Vermont in NIAAA-supported research, and discovered evidence of possible reactivity in connection with the use of daily telephone Interactive Voice Response (IVR) monitoring of regular alcohol users. In their study, consumption of alcohol decreased tangibly in association with the frequency of monitoring over a 2-year period, even in the absence of any specific intervention (Helzer, 2002; Searles, 2002). Specifically, drinking subjects received incentives to make daily telephone calls to answer a series of questions, lasting about two minutes total. Searles, Helzer & Walter (2000) describes the incentive schedule according to which a perfect calling record could net the subject approximately $13 per week.

The study evidence, originally collected for research on patterns of alcohol use among males, hinted that daily monitoring alone actually might decrease alcohol consumption. These findings prompted us to ask whether LSE-enabled daily monitoring might affect the target responses of major interest to us e.g., could mood be influenced by daily monitoring.

The concept of daily monitoring of mood is not unique to this study. The Life Chart Method was developed at the National Institute of Mental Health for

monitoring mood in individuals with Bipolar disorder. The Life Chart Method has been adapted for self-reporting both prospectively and retrospectively (Leverich and Post 1993). Although it is not unique as a measure of mood state on a daily basis, it does offer unique features for bridging the participant's behaviors and his or her mood state. The Life Chart Method Prospective Self-Report and Retrospective Self-Report Manuals provide keyword guides for individuals to rate their mood state based on changes in function and behavior (Leverich 1997a, 1997b). The Life Chart Method also provides a mechanism of self-report of behaviors that influence mood, such as substance use. Finally, the Life Chart Method allows the participant to rate life events that have impacted the individual on a scale ranging from positive to negative. These life events may influence the participant to engage in behaviors (drinking) that affect mood both directly and through subsequent behavioral changes, e.g. not going to class and failing. The Life Chart Method has been commonly used as paper and pencil form completed by the individual at the end of the day. The value of this daily mood monitoring has been well described in the literature for Bipolar disorder, but is it suitable for this study? This exploratory study is designed to lay a foundation for future studies of mood disorders treated with on-line interventions and assessed with on-line outcome measures via the LSE. In this approach, the Life Chart Method is a useful instrument. The ease of interface with the LSE facilitates participation and offers the opportunity to extend frequency of monitoring with instruments that already exist. We do not know, however, if the frequency of monitoring will affect the participant's mood.

SECTION2.J: Innovative biostatistical approaches

During the course of designing a trial to test for reactivity of daily monitoring, we came to appreciate a newly developed 'principal stratification' framework for drawing causal inference from evidence of experimental trials of this type. Specifically, we appreciate that there is likely to be some degree of non-compliance with any randomly assigned condition of monitoring. For example, when we randomly assign participants to daily online monitoring of mood, we cannot guarantee that every participant in this arm of the trial actually will log in, day after day, with fidelity. In addition, some participants assigned to the monthly monitoring condition will be conscientious and will record their daily mood state on a wall calendar, day by day, even though they have been assigned to the monthly online assessment schedule.

Under these circumstances, with non-compliance of both of these types, the proposed trial has an analogue in the concepts of the 'broken experiment,' one version of which was described recently by Barnard (2003). In specific, for that study, families were randomly assigned to receive vouchers to help finance a child's private school education, rather than having the child go to public school. As might be imagined, for some of the participating parents who were assigned to the 'no voucher' condition, the response was to send the child to private school anyway (e.g., self-financed without vouchers). Plus, for some of the parents who received vouchers, they sent their children to public school, despite having received the voucher for private school.

As outlined in a series of papers by Frangakis and colleagues, including Barnard, (1999) empirical application of the 'principal stratification' framework for causal inference can be tailored to 'broken experiments' of this type, with improved inferences about the causal effects of the interventions under study. Background papers by Frangakis, the lead biostatistician on our research team and a co-investigator, are available online for this proposal, including two theoretical papers (Frangakis & Rubin, 1999; 2002), as well as (Barnard, 2003; Frangakis, 2004). Please see links in references. Implication of causal effects in partially controlled studies continues to be the focus of research and relevant to this proposal (Li & Frangakis 2005 and 2006). Dr. Frangakis' effort within this field of study will guide our efforts in the analysis and interpretation of results.

In situations of this type, the new framework for causal inference is not analogous to 'off the shelf' software (e.g., procedures or commands under STATA, SAS). Rather, the framework must be tailored to the specific research application. Hence, one of the specific aims of the proposed research, to be completed as part of this project via an MSU-JHU subcontract to support Dr. Frangakis' work, is a specific tailoring of the principal stratification framework to the aims of this empirical study. As such, the result will be additional development and refinement of this important statistical methodology for future NIH trials – many of which qualify as 'broken' experiments of the type described above.

SECTION2.K: Rationale for starting with college students

This step in our research program begins with college-attending young people. As described in Chapter 2, the college student population makes up one forth of the population of 18-24 years in the United States. This population has ready access to the Internet and the ability to use the Internet. Within college students, one in ten reports having been diagnosed with depression. The public consideration of mood symptoms and depression in college students is not a sudden current concern. The National Institutes of Mental Health has made available public information concerning the presence and the impact of depression on college students for many years (http://www.nimh.nih.gov/publicat/students.cfm). The need as well as a means for this method to assess mood in the college students via the Internet offers a high potential for success of this proposal. In this research project, the study population might not be widely representative of the general population. It does represent an important population for interventions, particularly those delivered online. The PEW Internet and American Online Project has surveyed students and adults concerning their on-line behaviors since 2000 (http://www.pewInternet.org). The trend reported by PEW Internet and American Online Project among all individuals is a steady annual increase using on-line resources and engaging in on-line activities. Within the college age population, 72% go on-line daily to at least check email. The teen population is using the Internet daily at a rate of 51% up from 42% in 2000. The teen population has also shown a dramatic increase in getting health information on-line. Thirty-one percent of teens get health information on-line which is an increase of 47% since

2000. There is an upward shift in younger individuals using the Internet and mental health research can make use of this trend. The limitations of pre-Internet accessibility to subjects in research have begun to dissipate with the advent of the so many individuals having on-line access. The use of email and on-line access provides a mechanism to gather information. This rising trend is having an impact on research. Studies of medication compliance have used on-line resources such as email and on-line surveys. The likelihood of participants being able to complete instruments for assessing features such as side effect reports seems feasible if not even necessary for rapid reporting of problems. The risk of privacy issues and confidentiality in research online remains a serious concern. In order to determine if daily monitoring affects mood, the study will be conducted via an anonymous mechanism. The potential effect on mood due to daily online assessment needs to be delineated in the face of rising trends of online activities.

SECTION 2.L: Potential Public Health Significance

If successful, the proposed R01 research project will enhance our methodologies for a future research program to evaluate online health promotion stimuli (e.g., public service announcements via flash videos) and health promotion programs (e.g., more intensive multi-session programming via the Internet). The topic of the current R01 project is assessing the participant's mood state, but the methodologies and experiences to be developed will have pertinence across a broad range of depressed mood states.

The tailoring of the principal stratification framework for causal inference under conditions of incomplete compliance and 'broken experiments' can be used to illustrate this more general public health research significance of the proposed project. The tailoring of the principal stratification framework for this study will carry over to other evaluation research projects focused upon other lifestyle facets such as diet and exercise. For example, in evaluating an experimental online diet intervention program some of the participants assigned to the experimental arm of the study will not comply with diet recommendations, whereas some participants assigned to the control arm will engage in extra-protocol diet maneuvers similar to or with bio-equivalence to the diet maneuvers of the intervention arms. A similar 'broken experiment' character is presented in any evaluation of online interventions to improve regular exercise as a health-promoting behavior.

Finally, a major contribution of the proposed R01 research project is new evidence on a topic of potentially considerable importance for the design of future online intervention trials, as well as trials that apply online surveillance and monitoring methods to more traditional interventions. Namely, we seek evidence on the topic of the reactivity of daily monitoring, which often is needed if we are to gain fine-grained evidence about mechanisms and processes through which interventions are having their effects and the time-sequences of events and processes that lead toward success or failure of interventions.

Epidemiological surveillance by traditional methods typically yields relatively coarse-grained temporal sequencing data, often too coarse-grained to

shed light on mechanisms and processes of this type (e.g., see Wu & Anthony, 1999; 2000). The LSE methodology offers an opportunity to gather fine-grained evidence within balanced cost parameters (due to automation of the data gathering), but it will be important to learn the degree to which daily monitoring, enabled by the LSE method, might be inducing reactivity of the type found in Helzer's research on telephone monitoring of drinking behavior (Helzer 2002).

It is for these reasons that we have focused upon this project's specific aims and have assembled a research team with a breadth of skills and research interests. We are enthusiastic about the project's chances to break new and innovative ground in NIH-supported research, and we hope that the NIH study section will share in this enthusiasm for innovation.

CHAPTER 3: PRELIMINARY STUDIES

SECTION 3.A: The Research Team

Dr. Paul Quinlan has assembled a talented research team for completion of this R01 research project. He will serve as Principal Investigator, but James Anthony, Ph.D., an experienced NIDA investigator in epidemiology and experimental prevention research will join him in leadership of the project as Co-Principal Investigator. Other scientists on the research team include John Helzer, Professor at UVT, with whom Dr. Anthony has worked for many years, and Constantine Frangakis, Ph.D., Associate Professor at Johns Hopkins, who is the project's lead biostatistician and co-investigator responsible for tailoring the principal stratification framework for this specific project's needs. Dr. Frangakis and Dr. Anthony have previously worked together in an application of the principal stratification framework to NIH-sponsored tobacco prevention research now underway. Collectively, this research team has the combination of clinical, epidemiological, and biostatistical skills and experience to make the project successful. Dr. Quinlan and Dr. Rios' involvement in the project will be supported under this award. Dr. Anthony is supported under a NIDA K05 Senior Scientist award. Dr. Helzer will be supported as a consultant, and Dr. Frangakis is supported under an MSU-JHU subcontract with MSU financing, which will be sustained via an NIH-supported MSU-JHU subcontract should this project be awarded funding.

Paul Quinlan, D.O., will serve as Principal Investigator (PI) for this project. He is a Co-Investigator for the Youth-Nominated Support Team (YST) Intervention for Suicidal Adolescents. This randomized, open label, active control, parallel assignment, efficacy study seeks to identify effective treatments for suicidal adolescents. Dr. Quinlan's expertise includes monitoring the YST study participants' risk factors for suicide and initiating safety protocols to reduce suicide risk.

As Co-PI, Dr. Anthony is an experienced epidemiologist and intervention researcher who has had continuous research support via NIDA R01 project awards since 1986. His c.v. lists more than 250 publications on topics pertinent to psychiatric epidemiology and the epidemiology of drug dependence, including many peer reviewed articles on depression. ISI Thomson lists him as one of the top 200 most highly cited contributors to the research literature in two separate domains, Psychology/Psychiatry, and Social Sciences. He is the chief architect for the Longitudinal Study Engine development at MSU. His activities as Co-PI for this project will be supported in part by his active K05 Senior Scientist award from NIDA with the remainder through this grant.

Dr. Frangakis is an Associate Professor of Biostatistics at Johns Hopkins University Biostatistics Department in the Bloomberg School of Public Health. Dr. Frangakis is widely published in the area of developing designs and analyses to better evaluate treatments in medicine, public health and policy (causal inference). He is particularly interested in randomized studies in which subjects do not comply with the assigned treatments and drop out. In recent work, he has

shown that the "intention to treat" method, which has been widely used for these situations, is not suitable to estimate the "intention-to-treat" effects (Frangakis & Rubin, 1999). He has recently integrated this research with colleague Dr. Rubin in the framework of "principal stratification and effects" for expressing and estimating causal effects under varied conditions.

Dr. John Helzer, Professor of Psychiatry at University of Vermont, is a leading clinician-researcher for experimental trials of innovative automated methods to augment brief interventions directed toward reduced heavy drinking, pain control and other health outcomes. His studies of the potential reactivity of automated telephone monitoring in relation to heavy drinking are important forerunners for the present study. He has been consulting with the MSU research team on aspects of this research design for several years. His consultation for this project will be supported via the Consultation line item in this project's budget.

Dr. Glynda Moorer is the director of the MSU Olin Health Center and will participate in this study as a Co-Investigator. The Olin service provides both triage and health care services for the MSU student population. Dr. Moorer is a crucial member of the project team by insuring that student-participants who suffer from mood disturbances will be identified and triaged for appropriate 24/7 Olin mental health services.

Dr. Leigh White is a psychiatrist at the MSU Counseling Center (part of Olin) that provides mental health services to all MSU students. Dr. White's role is to enhance mental health services access for study participants with 'urgent care'

needs for services, without respect to the randomized arm of the reactivity experiment (i.e., equal access for participants without respect to randomized condition).

Dr. Christopher Giuliano, MSU Psychiatry Department's Clinical Director, who will be responsive to student-participants whose responses signify need for treatment but do not want to be seen or cannot be seen through MSU Student Health Services (Olin or Counseling Center). Dr. Giuliano provides an anonymous service; students can provide an LSE certificate without having to reveal identities, student ID, or health insurance information.

SECTION 3.B: Preliminary Studies

The collection of data via the Internet has been implemented across the World Wide Web but little has been discussed about preliminary studies for these applications. The following two sections summarize the study design suitable for use with longitudinal study engine technology and describe pilot data obtained with the longitudinal study engine relevant to this application.

SECTION 3.C: Longitudinal Study Engine

This proposed study will make use of a new and powerful research tool, the Longitudinal Study Engine (LSE), recently developed at Michigan State University under the direction of Dr. James C. Anthony. The LSE is an Internet-based survey engine that offers anonymity and privacy for exploring topics that otherwise might be too sensitive to effectively engage participants in scientifically

meaningful disclosures. The LSE is designed for large sample longitudinal, epidemiologic, prevention, and intervention research protocols. The LSE is currently being used for other studies of college-age young people's health and behavior. For example, one arm of the LSE research is measuring changes in physical activity patterns over the lifespan and the impact of activity on cardiovascular disease risk, incidence, and mortality. A longitudinal cohort study design will be initiated during college and followed throughout a lifespan.

SECTION 3.D: Pilot test of Longitudinal Study Engine and Incentives

As described above, Anthony's research group at MSU developed the LSE for anonymous longitudinal and experimental intervention research. Its capacities include standardized online assessments, with embedded streaming digital video and audio as needed to enhance the online session. In addition, the research team members have completed LSE studies designed to estimate participation levels in Internet-based surveys across a randomly assigned gradient of incentive values. Conducted in April-May 2005 and August 2005, these first studies focused on a group recruitment context for participation in (1) a survey of health and behavior in students; and (2) a survey of DSM-IV PTSD history, depression, and work-related trauma in emergency responders. Both studies used the LSE developed at MSU.

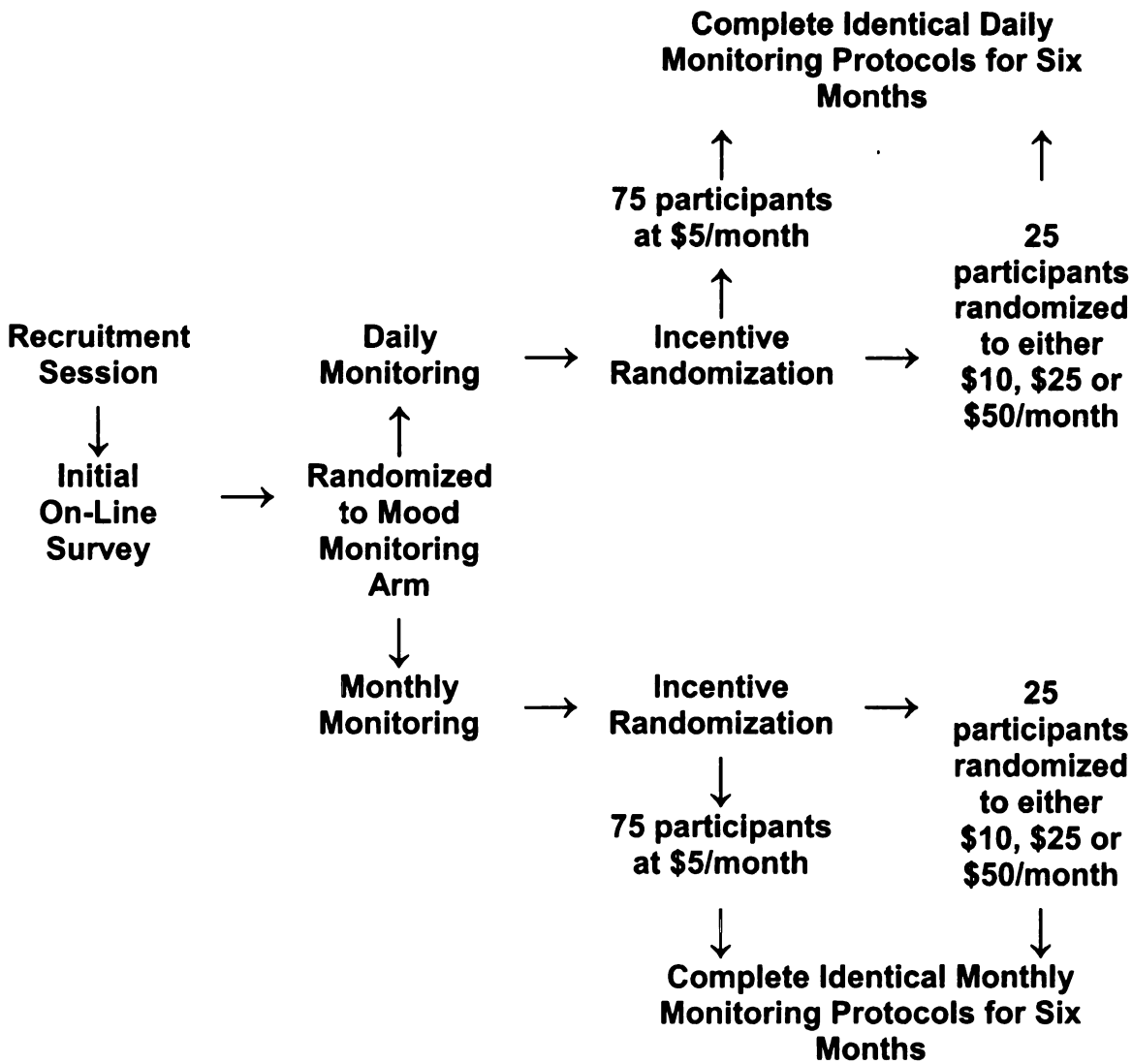
For the two separate studies, initial logins were compared as a function of level of incentive for 212 university dorm residents (Study I) and 148 professional firefighters (Study II). The initial login proportions for Study I were 33% of 76

dorm residents offered a $0 login incentive, 63% of 113 students offered incentives ranging from $2 - $18, and 91% of 23 students logged in among dorm residents offered $20 -$50, including 100% of the 11 students receiving incentive offers of $24 or more. In Study II, 29% of 51 firefighters receiving a $0 or $5 incentive logged in, 45% of 67 firefighters offered $10 logged in, and 90% (27 of 30) logged in among firefighters offered $25. On the basis of these studies, we propose a $25 incentive to encourage sustained participation in the proposed online survey research. Further information, please refer to www.epi.msu.edu/janthony under the title of "LSE Pilot Study."

CHAPTER 4: RESEARCH DESIGN AND METHODS

SECTION 4.A: Overview of Main Components in the Research Plan

Figure 1: Diagram of Participant Recruitment and Randomization



After probability sampling and recruitment from defined college-attending populations in residence on or near each participating university campus, an

expected epidemiologic sample of 4000-6000 18-25 year olds will be recruited for initial recruitment sessions in or near the potential participant's dwelling (e.g., dormitory floor lounge). At each initial recruitment session, all in attendance will draw a non-contingent prepaid cash value prize (value range, $0.25, $5.00), will listen to a study staff member read an IRB-approved introductory disclosure statement, and will be asked to complete a very brief anonymous optical scan survey form. All will be invited to log onto the LSE web site for more details about completion of an online survey of health and behavior, which will take no more than 8-10 minutes for completion. Results from the initial recruitment session op-scan form will permit study of initial non-participants versus participants who consent to visit the LSE web site and to log in anonymously for the online survey. (That is, all who agree to participate in the initial recruitment session will receive a prepaid prize of cash value in the range from $0.25 to $5.00; upon receipt of the prize, they will be encouraged but not required to complete the anonymous optical scan short-form survey.) Thereafter, via the LSE web-site, the consenting participants will become eligible for later randomly assigned cash value prizes designed to encourage persistence of log-in behavior and completion of the longitudinal trace of up to five online surveys, each requiring about 8-10 minutes of online survey time. Within this larger epidemiological sample, a projected 340-500 eligible regular participants identified via the initial online survey will be invited to participate in the experimental trial of daily versus monthly online monitoring of the participant's mood state, with incentives used to promote compliance with the monitoring assignment (e.g., variable interval cash value

prizes assigned at the completion of each daily online survey; fixed interval cash value prizes assigned at the completion of each monthly online survey). The principal stratification framework for causal inference will be used to investigate two main mechanisms of suspected causal effect: (1) the effect of randomization, and (2) the effect of daily monitoring versus monthly monitoring.

There are several important ingredients in this research plan:

(1) The MSU Longitudinal Study Engine (LSE), which has been designed to enable anonymous longitudinal surveillance of large epidemiological and clinical samples in health promotion research, with embedded randomized trials of online (and other) health promotion interventions, with probabilistic or algorithmic (deterministic) delivery of gratuity prizes to reinforce log-in behavior and survey completion behavior. As a longitudinal study authoring tool, the LSE gives the investigator control over important parameters for study design, including a broad menu of self-administration survey item response types (e.g., true/false radio buttons; Likert scales), as well as programmable schedules for participant log-in and delivery of tangible monetary or experiential reinforcers for participation (e.g., cash value prizes; flash videos of television comedy out-takes, etc.).

(2) Epidemiologic sampling and surveillance to identify participants who might qualify for recruitment into an online mood monitoring program. The participants will be recruited only from the student population at MSU.

(3) Periodic delivery of gratuities for participation, starting with delivery of a small tangible cash prize (i.e., a ticket redeemable for a stated cash value) at the time of initial recruitment. The designated respondent chooses the ticket envelope at random from an urn and learns the value of the cash prize as well as a unique ticket number that is linked to additional future gratuities for continued participation. This approach of prepaid delivery of a small tangible cash prize was chosen after a review of alternative prepaid and postpaid methods, in which the prepaid methods were found to yield higher survey participation levels. The periodic delivery of gratuities for subsequent participation is arranged in a mixed sequence of fixed interval schedules (at the time of initial log-in and monthly log-ins) and variable interval schedules (to reinforce frequent log-in behavior of the daily monitoring group).

(4) The ticket number, selected at random by each sampled participant, serves as a unique ID code that the participant records on all study materials. The ticket numbers are generated by the research team, but the research team does not know which individual has drawn which ticket numbers. After the ticket number is used by the participant to log into the Longitudinal Study Engine, the participant is invited to self-generate a unique userid, password, and memory prompts to recall the password. The ticket number is retained for tracking, along with the userid and password. Here also, each participant knows his or her own ticket number and self-generated userid, password, and password memory prompts,

but the research team does not, and has no way to link these values to any individual respondent.

The participants are offered the opportunity to be reminded to log-in on the assigned daily or monthly basis, via an automated email via the listserv mechanism, via an automated telephone calling system, or via a reminder sent by post or campus mail, but this option is managed by a third party independent of the research team (e.g., L-Soft, www.lsoft.com, or Majordomo). That is, the respondents who wish to exercise this option will be asked to mark a checkbox during the online survey session. At the close of the on-line session, they will be provided with instructions so that they can contact the third party that will activate the reminder system at the designated frequency (daily versus monthly). For example, those who choose the email reminder option will be given instructions about opting in (or out) as a recipient of the listserv messages sent by the commercial vendor, who is paid by the research project to maintain the listserv, receive the participant's email opt-in request (as well as any subsequent opt-out request), and generate the daily or monthly reminder according to the randomly assigned condition. The listserv-delivered email will include the URL for the LSE website so that the recipients can readily access the survey site, key in their individual user IDs and passwords, and complete the survey at that point in time.

(5) Random assignment of participants to the daily versus monthly monitoring arms of the randomized trial. The participants are being randomly assigned to

these two conditions in order to gauge the possibility that daily monitoring is a reactive measure with respect to mood, as compared to less frequent (i.e., monthly) monitoring.

(6) Recognition that the research team can randomly assign the condition of daily versus monthly monitoring, but cannot enforce compliance with these conditions. Rather, the research team can encourage compliance via the periodic delivery of gratuities as reinforcers or incentives, but some participants assigned to the daily monitoring condition will demonstrate non-compliant responses and will not engage in completely daily monitoring. In addition, some participants assigned to the monthly monitoring condition may be non-compliant, and in fact, some of them might self-initiate daily monitoring activities (e.g., writing down the change in mood over the day, using a wall calendar to record the values). This recognition helps to motivate elements of the principal stratification framework for causal inference under these conditions.

SECTION 4.B: Overview of the Research Plan'S Innovative Aspects

A goal is to refine the anonymous LSE daily monitoring plan so that daily monitoring might be used in future randomized trials to assess the impact of online cognitive behavioral therapy for depression programs, but to do so permitting an experimental test of the reactivity of daily monitoring versus less frequent monthly monitoring. The protocol for the proposed study's monthly online monitoring group is identical to the daily monitored arm, except that the

frequency of online monitoring is allowed to be monthly. The participants in the monthly monitoring arm are not allowed to log in on a daily basis. Otherwise, we seek equivalence and balanced distributions for the two groups of participants, with daily versus monthly monitoring as the key experimenter-manipulated contrast. To facilitate the balance, the participants will be cautioned not to share with other individuals their responses to the questions in the study. In addition, the participants also will be cautioned to not record their answers outside of the LSE system.

An innovative variation of this proposed research plan with respect to almost all prior mood assessment research is that we seek to investigate two main mechanisms involved in the study. One mechanism involves the effect that randomization has on receipt of the intervention, and the other mechanism involves the effect that receiving the intervention has on the outcomes. As explained above, these two mechanisms come into play because we cannot enforce 100% compliance with either the daily monitoring assignment or the monthly monitoring assignment. As explained below, in our analysis plan, conventional approaches such as 'intent to treat' approaches can be sub-optimal with respect to estimation of these effects. It is for this reason that we have turned to the principal stratification framework for causal inference.

The value of the principal stratification framework can be enhanced to the degree that mechanisms that give rise to non-ignorable missing data have been investigated thoroughly. For this reason, we have proposed a second innovative variation with respect to almost all prior mood assessment research, which is to

specify the population under study in advance, and to use randomly assigned tangible gratuities, incentives, and reinforcers to help gain experimental control over (a) initial consent to participate, and over (b) the online log-in behaviors required to be compliant with the daily versus monthly monitoring.

A third innovative variation from the conventional design is in our attempt to provide the participants with complete anonymity.

SECTION 4.C: The Study Population and Sample

To begin, the study population will be specified as community-dwelling college-age young people, mainly 18-25 years of age, drawn from multiple jurisdictions inside and outside the United States, attending college/university, and living in dormitories and other residential units on or adjacent to the participating university campuses. The sampling frame will be provided by the university administration as a list of all such dormitories and residential units, and a probability sample of these units will be drawn; all residents within sampled units will be invited to participate. Student assistants on the research team will be trained to recruit and roster the residents of each sampled unit and to read disclosure and informed consent scripts for recruitment, as approved by the cognizant institutional review board (IRB).

In total, the study population will consist of slightly more than 30,000 residents, many of whom will be living in multi-person group quarters such as dormitory floors, with 50 residents per floor. This grouping of sampled participants within residential units will facilitate completion of group sessions for

the initial recruitment session. For this study, the study population will be sampled to yield 3500-6000 probabilistically designated respondents 18-25 years of age, each of whom will be invited to the initial recruitment session(s) described below. The LSE pilot study now underway will help the research team to calibrate the sampling fractions and sample size within the range from a minimum of 4000 designated respondents to a maximum of 5000 designated respondents; the budget has been specified to allow up to 6000 designated respondents. Based on prior online surveys of the local college populations, the expectation for online survey response levels is 20%, with no prizes or incentives/reinforcers (i.e., 1500 of 6000). With the planned incentives/reinforcers, the expectation increases, and proves to be sufficient to generate the sample size of eligible participants as specified below.

SECTION 4.D: Participants and the Initial Recruitment Session

All sampled participants, after initial consent, will receive a prepaid cash value gratuity as an incentive to participate, with the cash value constrained to be small (range = $0.25, $5, with variations in units of 25 cents), and self-drawn at random by the potential participant using a lottery method. That is, the potential participant will draw the prize at random and will know the value of the prize before deciding whether to participate.

The value of this initial small but tangible cash value prize will be printed on a prize drawing ticket. Also printed on the prize drawing ticket will be a multi-digit ticket number that serves as a unique ID code for the individual participant,

as well as the survey website URL, and the value of an additional cash value prize that the participant will receive when the subsequent log-in action occurs. Typically, the login will be scheduled to occur within one week of sampling and recruitment for participation.

If they will consent to do so, at the time of this initial designation for the sample and the initial consent, all participants will complete a brief optical scan survey form to record the following details: self-rated general health status, smoking status and frequency of online Internet activity, sex, age in years, and minority status sufficient to provide NIH with required reporting about participants. They will be asked to copy the ticket number onto the optical scan survey form, and will be told the reason for gathering this information, which is to permit a comparison of those who actually log-in for the online survey versus those who consent to take the online survey but who fail to do so.

The initial recruitment session will end at this point, but the student assistant may return to the residential unit up to 10 times in order to seek participation and recruitment of sampled unit residents who are not present or available for recruitment at the time of the initial visit and session.

SECTION 4.E: Initial Online Session: Log-In & Recruitment

As described above, the prize drawing ticket will list the survey web site URL, a unique ticket number, and the value of an additional prize to be gained simply by logging onto the web-site and reading the IRB-approved description of the epidemiological surveillance approach. The survey response level will be

controlled, in part, by the value of the additional prize, which will be range in value from $0 to $50, with at least 50% of the values equaling $0 (i.e., no tangible prize), and with the vast majority of the values constrained to be under $5. Nonetheless, the probability of drawing a prize of cash value $20 will be at least 1 in 100 and the probability of drawing a prize of cash value $50 will be at least 1 in 500. Participants will be advised of these probabilities at the time of the initial recruitment session.

Upon accessing the survey web-site via its URL, the LSE participant is instructed to key in the unique ticket number, to self-generate a user ID, password, and password memory prompt, and to read (with audio accompaniment) a brief disclosure statement that describes the terms of anonymous longitudinal surveillance as described below. Thereafter, the participant is credited with the prize and can print a redemption coupon at that time, terminating future participation. Alternately, upon reading the on-screen text and listening to the audio recording of the IRB-approved survey disclosure statement, the participant can become eligible for additional prizes by clicking a check-off box that signifies "I agree to participate in this online survey," and by completing the initial online survey.

SECTION 4.F: Initial Assessment for Anonymous Surveillance

As just noted in Sections D2.2 and D2.3, a potential participant can opt out of participation at the time of the recruitment session, or after the initial log-in session. The ticket number on the optical scan short form will permit

characterization of participants who log in versus those who do not (e.g., with respect to demographics, etc.).

For those who consent to complete the initial assessment as part of our anonymous longitudinal surveillance, there will be an online survey of roughly 8-10 minutes duration, with introductory standardized survey items on self-rated general health status and other health-related topics that we have found to promote trust and rapport during early parts of health interview studies (e.g., in the NIMH Epidemiologic Catchment Area Surveys of 1981-85, and in other subsequent research on adolescent and young adult drug involvement; e.g., see Storr 2004).

The content of this initial online survey assessment is specified below in Table 1.

Table 1: Survey Assessments

Survey	Subjects	Data collected	When
Recruitment Session	All DR = designated respondents	Very short op-scan survey form to assess mood state and Internet usage, socio-demographics required for NIH inclusion tables.	During a recruitment session after introduction by staff.
Initial online survey	All DR are invited to log in for the initial online survey.	Demographic, background characteristics; history of mood symptoms; concomitant use of alcohol or other drugs; interest in Internet programs to change behavior, study habits, grade point average.	First survey for entry into LSE study only
Baseline assessment for the trial	Eligible participants	Baseline criteria provide information for the study and evaluation for possible factors that could confound the study such as concurrent psychotropic medication and/or therapy for mood symptoms.	If convenient, these items are completed after consent for trial participation, either during the initial online survey session or at a baseline time set by the eligible participant.
Daily survey	Eligible Participants	In reference to the past month: Any outside resources sought for mental health	Offered monthly for 6 months
Monthly survey	Eligible Participants Assigned to Monthly Online Monitoring Arm	In reference to the past month: An assessment is made across the participant's mood state for the month by the NIMH Life Chart Method self-report version/retrospective and any outside resources sought for mental health	Monthly for 6 months (total of 6 surveys including baseline)
Daily survey	Eligible Participants Assigned to Daily Online Monitoring Arm	In reference to the past day (one time per day only): An assessment is made across the participant's mood state by the NIMH Life Chart Method self-report version/prospective	Daily for 6 months, with monthly assessment in place of daily assessment. Content of the monthly assessment is that same as is described above for all eligible participants assigned to monthly arm.

During completion of this general health and health-related behavior survey, the automated algorithms of the LSE will be used to lead eligible participants to a branch of the survey that involves recruitment for the experimental trial of daily versus monthly monitoring. This branch will present the eligible participants with an IRB-approved disclosure statement that describes the experimental trial of daily versus monthly monitoring of mood as described in Section 4.G, which involves completion of the trial's baseline assessment, and then log-ins at the assigned frequency (daily versus monthly), with periodic delivery of gratuities to serve as tangible reinforcers of the log-in behavior, as described in subsequent sections.

SECTION 4.G: Recruitment and the Experimental Trial Assessment

Based on our online analyses of recently gathered national surveys, among 4000 probabilitistically sampled college-attending young people, a minimum of 700 of the online survey respondents would qualify for this eligibility requirement and would not be excluded by virtue of current therapy for depressive symptoms. (A projected 880 would qualify prior to the online survey.) Because this is an exploratory research project, that projection could be off by a factor of 50%, and the project still will yield valuable results (i.e., with only 350 meeting eligibility criteria).

During the initial online session, the LSE branching algorithm will lead all of the eligible participants to the IRB-approved disclosure statement that describes the daily/monthly online monitoring trial and elicits formal consent via a

marked checkbox. Eligible participants who mark the checkbox will be instructed to continue the online survey session at that time, in order to complete an additional 10-20 minute baseline assessment for the trial, or to log out of the system until they can log back in to complete the additional 10-20 minute baseline assessment for the trial at a time more convenient for the respondent.

The IRB-approved disclosure statement will explain the randomized assignment to the two arms of the trial, daily online monitoring versus monthly online monitoring, as well as the reinforcement contingencies outlined in Table 2.

Table 2: Incentives/Reinforcers for Survey Completion

Survey	Incentive value per survey completed
Recruitment Session (all will be asked to complete brief op-scan form; cash value ticket is non-contingent; i.e., op-scan form not required.	Prepaid cash value ticket drawn at random by the potential participant at the start of the recruitment session (value range = $0.25, $5, with variations in units of 25 cents)
Initial On-Line Survey	50 students each will receive $10, $25, or $50; rest will receive a value in a range from $1.00 to $10.00 (mean ~$4)
Eligible Participants in Monthly Monitoring Arm	Set amount for all 6 subsequent surveys at monthly intervals: 25 students each will receive $10, $25, or $50; rest will receive $5
Eligible Participants in Daily Monitoring Arm	Set amount for all 6 subsequent surveys at monthly intervals - 25 students each will receive $10, $25, or $50; rest will receive $5. For daily online surveys, a variable interval schedule will be used, with delivery of cash value incentives in a range from $0.25 to $2.00. All are eligible for $3/week bonus when all seven surveys are completed in a seven day interval.

The content of this baseline assessment for the daily/monthly monitoring trial is described earlier in the final rows of Table 1:

Mini-CES-D: The CES-D or Center for Epidemiological Studies – This Depression scale is a commonly utilized measure for depression screening

(Radloff, 1977). This scale has been used reliably in examining the relationship between depression and Internet use (Morgan & Cotton, 2003), both in college students. A short form of this tool has been tested, with resulting evidence supportive of its validity (Cole, 2004). A study by Wight, Sepulveda, & Aneshensel (2004) found in a study of adolescents to adults on the CES-D that older adolescents reported the highest depression values, but persistence of clinical features was similar between adolescents and young to middle aged adults.

AUDIT: Alcohol use often co-occurs with depression and might be a determinant of mood. Therefore, the examination of alcohol use and patterns is important to this study. The alcohol use disorders identification test (AUDIT) was developed by the World Health Organization to identify persons who alcohol consumption has become hazardous or harmful to their health (Babor 2001). The AUDIT is a 10-item screening questionnaire with 3 questions on the amount and frequency of drinking, 3 questions on alcohol dependence and 4 questions on problems caused by alcohol (Babor 2001; Saunders 1993). Numerous studies of validity and reliability have been conducted (Allen, 1997; Volk, 1997) and this measure is generally accepted to be valid and reliable for use with college students (Allen 1997; Volk 1997).

Drug Use: Use of drugs other than alcohol and tobacco also co-occurs with depression, and might serve as a determinant of mood. For this project's anonymous online survey assessments of health and behavior, we will draw from the key illegal and extra-medical drug involvement questions that are part of the

Monitoring the Future (MTF) surveys of college students (Johnston, 2004). Based on MTF results, we have identified the most commonly used drugs among college students and have included questions on those drugs only, as to minimize the length of our survey.

SECTION 4.H: Ongoing Assessment of Mood

The NIMH-Life Chart Method-TM (LCM™) Self-Version: The assessment of mood for this study will require prospective and retrospective reporting. Hörn has described long-term monitoring methods in mood disorders (Hörn 2002). Of those methods described, the NIMH Life Chart Method offered advantages in validity, self-completion versions, clinical relevance and ease of availability (Denicoff 1997, 2000, 2002). The monograph by Leverich (Leverich 1997a, 1997b-see references for hyperlink) describes the approach for the use of the self-report version for both the prospective and retrospective forms. Although these self-assessment instruments are designed for individuals with bipolar disorder, there are two reasons for consideration of their use in this study. The first is the relative ease of use. The keywords provided by Leverich (Leverich 1997a and 1997b) guide the participant in rating his or her mood. The scale includes both elevated and depressed descriptors that are critical to assessing mood reactivity to on-line monitoring. The mood rating is a single score reducing the time and keystrokes needed to complete the assessment. The second consideration for using the LCM™ is that this is an exploratory study that will determine if on-line monitoring will be feasible for studies of affective disorders.

The outcome of this study would facilitate the study of mood disorders via the LSE. The LCM™ had been used in studies of affective disorders and is a well-established instrument for tracking mood. It has also been converted into a Palm handheld electronic form. The linkage of the LCM™ to the LSE will unite the strength of close monitoring of mood state and the sophistication of the LSE in its ease of use in collecting data.

SECTION 4.I: Access to Mental Health Services for Participants

For all study participants, a health services information page about access to the Olin Health Service and its MSU Counseling Service will be included in the orientation login session when the anonymous participant connects to the LSE server for the first time. This information page consists of the same information provided by MSU to incoming students at orientation. This page will review the services available to the student participant and how to access these services.

In the event that during a session any student participant indicates a severe symptom on the NIMH Life Chart Method scale, he or she will be automatically advanced to a mental health help page sent by the LSE server during the current online session. The mental health help page will invite the participant seek to contact the Olin Health Service. The page informs the participant that he or she can contact the Olin Health Service to speak with a nurse on call about his or her symptoms and options for help. The Olin emergency telephone numbers will be listed on the page. The Olin Health Service provides all MSU students with 24 hour a day telephone access to Olin

nursing staff who triage health and mental health questions from students directing to the appropriate service. This service will also accommodate calls from participants in this study. The nurse will triage the study participant for risk behavior to insure that proper treatment is initiated at the time of the call. This assessment follows the same assessment algorithms currently in use for triaging students through Olin Health Service. The participant does not need to identify him or herself as a participant in the study to receive assistance from the nurse and will be informed of this on the mental health help page. An anonymous notification that a participant has identified severe symptoms will be sent to the PI. The LSE survey will include a follow-up question page at the next login session for the study participant who was notified to contact Olin Health Services due to severe symptoms recorded on the NIMH-LCM. The follow-up question page will ask the participant if she or he did call Olin Health Services or 911. The possible responses to this query for the participant to select from are "Yes and I did receive help", "Yes, but I did not receive help" or "No, I did not call". In the first response the participant will be advanced to the appropriate rating scale for the study. If the participant chooses the second response, she or he will be offered the option to seek further assistance. The participant will be offered the same numbers again as well as offered the telephone number to request services through MSU Department Psychiatry. The participant will be given an office telephone number for Dr. Giuliano, a clinical psychologist and MSU Outpatient Psychiatry Clinic Director. The MSU Psychiatry Clinic option is provided to accommodate a participant who began the study as a student but

may no longer be attending MSU. The MSU Psychiatry Clinic option also offers an option for further assessment and treatment for the student who does not want to be seen in MSU Student Health Services. This step provides two options for a participant to speak to a mental health professional for assistance in the event severe symptoms occur during study. The student participants who reach this level will be given an alternate follow-up question page at the next login to determine outcome. If the student participant still has not received help, he or she will be prompted to contact the PI for further assistance.

Dr. Moorer, Dr. Giuliano and Dr. White will meet monthly with Dr. Quinlan to review the frequency of anonymous student participants prompting referral to Olin Health Services or MSU Psychiatry. The clinicians are aware that these students seeking help are anonymous participants in the study and participant anonymity will be maintained in the meetings with Dr. Quinlan.

SECTION 4.J: Statistical Analysis

The following sections describe the principal stratification framework critical to this study. Also discussed are the methods for data management and sample size calculation.

SECTION 4.K: Principal Stratification Framework

To evaluate the hypothesized effect of daily online monitoring relative to monthly online monitoring, and to provide information that can be used to improve mood interventions for different segments of the population, it is

important to estimate two main mechanisms involved in the study: the effect that randomization has on receiving the intervention of daily online monitoring.

A conventional approach to analysis of data from a trial of this type might include elements such as: (a) discarding information about the base population from which the trial participants were recruited (e.g., ignoring information about participants who declined to give informed consent for participation); (b) focusing the analysis upon those consenting to participate and assigned at random to the 'daily' versus 'monthly' monitoring arms of the study; (c) ignoring the facts that some participants assigned to the study-devised 'daily' monitoring group might actually engage in self-devised monthly monitoring and that some participants assigned to the study-devised 'monthly' monitoring group might actually engage in self-devised daily monitoring (e.g., via a daily exercise of writing down their mood state on a calendar), as opposed to the study-devised on-line monitoring method. Most of above conventional approaches are problematic because they compare non-comparable groups. For example, individuals receiving the intervention can be different from those refusing it, in terms of underlying characteristics related to the outcomes, because compliance with the intervention is not controlled but selected by the individuals. Then, a comparison between the groups receiving and not receiving the intervention does not estimate the effect of receiving the intervention.

For these reasons, we deliberately chose to abandon these conventional approaches in favor of a methodology that seeks to estimate the effect of daily versus monthly online monitoring under explicit assumptions using the framework

of principal stratification for causal inference. Frangakis and Rubin (1999, 2002) describe alternative frameworks for causal inference, and refine the principal stratification framework as an elaboration of Neyman's early 20th century concept of 'potential outcomes' for inference from the evidence of randomized experiments, and with building blocks in the form of Rubin's elements for causal modeling such as 'ignorable assignment,' 'propensity scores,' and 'sequential ignorability.' The framework of principal stratification for causal inference encompasses and extends these elements; its recent application includes a series of empirical studies on a variety of topics, including (i) the effect of school choice vouchers and public/private schooling on children's school achievement (measured via standardized test scores), as estimated under conditions of a 'broken' randomized experiment in which some parents who received randomly assigned vouchers for private school did not send their children to private schools and some parents who were randomly assigned to the 'no voucher' condition actually sent their children to private schools (e.g., see Barnard, Frangakis, Hill 2003), and (ii) the effect of a Needle Exchange Program (NEP) on risk of and time to HIV sero-conversion (Frangakis 2003).

In its application to this project, the principal stratification framework is well suited to estimate the effect of the intervention on main study outcomes such as alteration in mood state. Specifically, a principal stratum of an individual will be that individual's mood state. The principal stratum exists because it is defined based on the controlled level of incentive, but is not entirely observed: at a specific time, only one of the individual's compliance behaviors can be observed,

the behavior for the incentive to which the individual will be actually assigned. Two questions then are relevant here: why is principal stratification important? And, how can we estimate quantities related to it?

First, principal stratification is fundamentally important in defining the effect of receiving versus not receiving the intervention, even though receipt is selected by the individuals. In particular, it has been shown that we can express a comparison of outcomes under receipt and no receipt of the intervention for exactly comparable individuals as an experimental comparison of outcomes between different levels of incentives if we condition that comparison in certain principal strata (Frangakis and Rubin, 2002). Second, although the principal strata are partially unobserved, we can statistically estimate this latter comparison under assumptions that are generally more plausible than those made by standard methods. More detailed coverage and motivating examples are provided in the published articles (e.g., Frangakis & Rubin 1999, 2002; Barnard 2003; Frangakis 2003).

One point of departure for the explanation is the now-fairly-conventional standard 'Intention to Treat' (ITT) approach, which we will use to derive initial estimates of the daily monitoring effect on the response(s) of interest here. That is, the ITT contrast expresses the depression level for those assigned at random to daily monitoring versus those assigned at random to non-daily monitoring – even when some of those assigned to daily monitoring actually do not comply with the daily requirements of the daily monitoring condition. However, as demonstrated by Mock (2005), additional especially useful information about

intervention effects can be gained via the PS approach. In brief, the PS approach starts with estimation of the complier average causal effect [CACE] under the dual assumptions of monotonicity and exclusion as described by Angrist (1996). The monotonicity assumption should not be terribly controversial in this context; it implies that any participant who would experience a changed depression level when assigned to the control condition (non-daily monitoring) also would experience a changed depression level if the random assignment had been to the experimental intervention condition (daily monitoring). The exclusion assumption implies that any ITT effect of assignment on the response of interest is mediated by assignment's effect on intervention received. As noted by Jo (2002), violations of the exclusion (restriction) assumption can be accommodated with careful model inclusion and measurement of covariates that predict compliance (e.g., expected compliance as measured via standardized LSE items we already have described, such as 'How likely or unlikely is it that you would participate in daily monitoring of depressed mood if you were offered the opportunity to do so during the next week?'). This exclusion assumption implies that there is no effect of randomization on level of depressed mood if the participant qualifies as an 'always-taker' or 'never-taker' (i.e., no matter what the randomized assignment might be). Under the monotonicity and exclusion restriction assumptions, CACE can be derived asymptotically (Angrist 1996) or in smaller samples with modeling of covariates as in Mock (2005) via unrestricted polytomous regression for the principal strata given covariates; the modeling is based upon maximum likelihood with a common variance normal model for the

responses of interest given PS and covariates. Sensitivity to different covariate specifications can be modeled, and a 'testimation' approach involves a two-sided approach with specifications for 5% Type I error and 95% nominal coverage for the confidence intervals.

We propose to estimate proportions of never-takers, always-takers, and full-compliers, as well as the mean response level for never-takers, the mean response level for always-takers, and the mean response levels for full compliers assigned (i) non-daily monitoring and (ii) daily-monitoring, with the difference of these two means as an estimate of the effect of assignment for full compliers, with the use of informative covariates to improve estimation of this difference. These covariates are summarized as an index I, representing a weighted sum of each participant's covariate values, the value of which has been demonstrated in work by others, with weights based upon the regression coefficients of the covariates in ITT regression that ignores assignment status (e.g., Cochran, 1968; Rosenbaum & Rubin, 1984). As such, the I indices for each participant stand for a propensity summary with respect to the response of interest if assignment is to non-daily monitoring, which is unaffected by assignment. As illustrated in Mock (2005), the participants can be sorted by quantiles of the I (propensity score) distribution, and indicators for membership in these categories can be introduced as covariates in the model, with one category left out to serve as a reference sub-category as in standard multiple regression modeling. The sensitivity of results derived via the PS method can be probed in relation to different covariate

specifications (e.g., via four and six sub-classes based upon quantiles of the l propensity index), as in Mock (2005).

SECTION 4.L: Data management and quality control

The LSE automated approach to online surveys is one that assembles the collected data within an established database format (e.g., Excel), and accumulates the longitudinal data over time within that format. The use of pre-specified response categories and standard online survey methods is an approach that reduces data cleaning problems, even when respondents deliberately skip certain questions (e.g., questions too sensitive for them to answer). The LSE approach addresses skipped items by asking the respondents whether they intended to skip the item, and by allowing them to skip or to enter an item response if they had not intended to skip. The LSE also automates creation of a codebook that lists the survey items and all possible pre-coded response values. String or character entries are allowed in short answer or extended essay formats. These values are encoded as text or string variable values within the LSE database.

In practice, our LSE team exports the Excel or other data to MS Access, Stata, SAS, or other software for statistical analysis, and conducts standard range and logic checks as part of the data management and cleaning steps. In addition, actual empirical frequency distributions are produced for each study variable with standard missing value conventions (e.g., for skipped items). When needed, editors and coders on the LSE research team can use ZyIndex or other

software to aid encoding of text and string responses.

Exploratory data analyses (e.g., stem and leaf plots) are used, as appropriate, in order to complete the data cleaning process and to enhance data quality prior to statistical analysis.

SECTION 4.M: Sample size projections

The principal stratification framework for causal inference is focused upon estimation of effects, and can be implemented with a perspective that is a departure from the conventional frequent position of testing departures against the null. Nonetheless, a review of this proposal might be interested to know that with no correction for nested structures or clustering of respondents within the survey design, a total trial sample size of 400 eligible participants would yield 80% power to detect a twofold difference in mood state for daily versus monthly monitored participants, with alpha set at 0.05. The survey design correction shifts the detectable relative risk parameter to the right, but even if the effective sample size is reduced to n=200 (100 in each arm of the study), the trial should detect a relatively modest or moderate size effect of daily monitoring.

Nonetheless, we are hopeful that the peer reviewers and study section will not place too much emphasis upon the issue of statistical power, especially in light of our application of the innovative principal stratification framework as described below.

SECTION 4.N: Scientific Writing

The research team expects to produce substantive findings and methodological advances of considerable interest to epidemiologists, clinical researchers, and others interested in the refinement of methodologies for large-sample randomized trials, anonymous online longitudinal surveys, and the substantive domain of psychiatric epidemiology.

The addition of tailoring the principal stratification framework to this specific research problem will also be of interest to biostatistical methodologists and others.

In order to ensure the dissemination of these results and methodological advances, our timetable includes an interval of post-data-collection work that will allow us to complete the statistical analyses, discuss and draw suitable inferences, and prepare scientific articles for publication.

SECTION 4.O: Proposed Project Timeline

The timeline for the project is described below in Table 3.

Table 3: Proposed Project Timeline

Activities	Pre Grant	Year 1				Year 2			
	4th Qtr	1	2	3	4	1	2	3	4
Phase I: Project Preparation: Monitoring intervention refinement, Instrument Testing									
Preliminary IRB approval at MSU	X								
Refine and test survey and assessment instruments		X	X						
IRB instrument revision approvals			X						
Phase II: Subject Recruitment and Daily Monitoring									
Send out information for initial survey				X					
Completion of initial survey				X					
Quality Assurance				X	X	X	X		
Phase III: Data Collection									
Daily/monthly monitoring				X	X	X			
Data capture				X	X	X			
Phase IV: Final Analysis and Report									
Data cleaning and analysis				X	X	X	X	X	
Write up of final reports							X	X	X
Completion and submission of manuscripts									X

SECTION 4.P: Anticipated Limitations of This Study

Q: What will be the representativeness of the population under study, and where will the colleges and universities be located?

A. For this R01 proposal we do not lay claim to national representativeness of the study's epidemiological sample. We can complete this study at MSU alone, given that our undergraduate student body consists of more than 30,000 college-

attending young people in the specified age ranges.

Q: How will the investigators prevent multiple survey completion by single individuals to get the incentives?

A: We have learned that from previous online survey experience that a very small proportion of these incidents should be minimal, in part due to the small mean and median value of the incentives. However, in addition, we embed non-identifying standard questions within each survey, with reference to stable characteristics (e.g., birth month of a respondent's mother), and we repeat these questions so that any instability generates an LSE signal of possible respondent fraud. We also discuss the issue of trust, validity of study evidence, and fraud as explicit topics in our recruitment sessions or in the initial online description of studies. As such, we are taking due precautions to reduce fraud of this type.

Q: Is there an issue with the incentives rates on a monthly basis being different for the monthly versus the daily monitoring groups?

A: The research team discussed the merits of providing the same amount of incentive per month for both monthly and daily monitoring groups. However, the burden of daily response is so much greater than the monthly response that it seemed inappropriate to calibrate the incentive values in an manner that would give the monthly monitored group as much as $70-$80 for completion of a single online survey, when the other group's operant responding requires daily log-ins and responses to secure that reinforcement value. Nonetheless, given that there

is a variability to the incentives/reinforcers distributed to the respondents, by virtue of the design of the study, there will be some overlap in the cumulative cash value of reinforcers assigned to each of these two groups. For example, the 25 students who will receive $50 for the monthly survey will be receiving a similar amount to the individuals completing the daily survey in the higher range of reinforcers (range from $26 - $68/month). This overlap makes it possible for us to calibrate these values in the middle of the distribution of cumulative reinforcers, and we can study effects, holding cumulative incidence constant.

Here, again, because this is an R01 exploratory/developmental research project, and because this issue comes part and parcel with aspects of innovation in this research design, we trust that the enthusiasm of our peer reviewers and study section members will not be dampened excessively by this thorny detail of methodology.

Q: With the anonymous sampling is there a potential for cross-over of the daily monitoring group into the monthly monitoring?

A: The longitudinal study engine (LSE) prompts the user to create a specific user ID code and password. The user ID then becomes linked with a specific set of surveys for that individual. Even if the person's identity is unknown, it is still possible to link the completed surveys to the same created user ID. As long as the programming is completed correctly, there is virtually no chance that a respondent would receive the wrong survey.

CHAPTER 5: HUMAN SUBJECTS

SECTION 5.A: Human subjects approval procedures

The primary institutional review board (IRB) for the participation of human subjects in this project will be the Michigan State University Committee for Research Involving Human Subjects (UCRIHS). To the extent that other colleges and universities join in the consortium arrangement and allow access to their own populations for college-attending young people, additional secondary institutional review boards may be involved (unless they allow the UCRIHS at MSU to function on their behalf.) The investigators believe this study presents no more than minimal risk to participating subjects, given the anonymous character of the Longitudinal Study Engine assessment. Nonetheless, there will be no contact with human subjects until the cognizant IRB(s) have granted approval to proceed.

To properly protect the rights of the subjects, the following procedures will be taken to minimize risks, ensure adequate disclosure and provide consent to the research and protected health information, and allow for voluntary participation. The research investigators and staff have been, or will be, certified in use of human subjects protections by the commencement of the project. MSU UCRIHS offers such an online course, upon completion of which the individual receives a certificate of completion for his/her records. All research staff will be required to participate and produce this certificate.

SECTION 5.B: Recruitment and consent procedures

As described above in section 4.G the recruitment will occur via research assistants (RAs), generally medical and nursing students employed or in field research placements with this project. All RAs will be required to review and sign a statement of understanding regarding their role in the project. This statement includes a description of the: tasks he/she is expected to complete to adhere to the research protocols, incentives provided for completion of these tasks, and expected ethical obligations regarding distribution of the admission tickets and confidentiality of information obtained through participation in this project. Resident assistants will receive human subjects use training as deemed necessary by the university institutional review board.

We will ask for a waiver of formal signed consent because the signature would constitute a disclosure of a specific participant's identity, and thereby creates an otherwise unnecessary linkage between the truly anonymous respondent and the study responses and outcomes. In lieu of signed consent, there is a series of online consent forms. Subjects who opt to log into the system will review the online consent form and will proceed with participate when they consent to participate. Otherwise, they will log out of the system. The IRB-approved disclosure and consent form will describe the study purpose, procedures, voluntary nature of participation, confidentiality/anonymity protections, an analysis of potential risk and benefits, and contact information of research officials (included in this section). As this study does not collect

protected health information as defined by the Health Information Portability and Privacy Act of 2001 (HIPAA), a request for HIPAA waiver will not be requested.

SECTION 5.C: Confidentiality/Anonymity procedures

The participants will remain anonymous. For example, there is an initial op-scan survey form that will be completed by the potential participants at the time of an initial recruitment session, but this form will not be used to record the potential subject's name, address, or other linkable identifiers. Rather, it will be restricted to the information required to fill out the NIH inclusion forms, and to information about matters such as frequency of recent Internet usage and description of the participant's mood state. We will not ask about illegal or 'sensitive' behaviors for this op-scan survey form, which is mainly to be used to characterize participants who log into the online survey site versus those who do not log in.

At the time of the initial recruitment session, all designated respondents will be given a specific admission ticket number that they will use to log on for completion of the initial online survey form, if they chose to participate.

The initial online survey and all subsequent surveys will be administered using the Longitudinal Study Engine (LSE), recently developed at Michigan State University under the direction of Dr. James C. Anthony, a Co-PI for this project. The LSE is an Internet-based survey engine that offers anonymity to longitudinal survey participants. The LSE is designed for large sample longitudinal, epidemiologic, prevention, and intervention research protocols. The LSE

procedure is devised so that an anonymous participant can log in, self-generate a personal userid and password, and remain anonymous to the research team (e.g., we do not ask for email addresses).

The initial online survey will present the IRB-approved disclosure statement and a checkbox to be marked if the participant agrees to become a longitudinal study participant. For any in the series of online surveys, the subject must continue to use the self-generated userid and password. The data files will never contain any identifiers that can link the information to the subject.

As explained in Section 4.E, all participants who log in and self-generate their personal userid and password will be eligible to complete a series of online surveys at either daily or monthly intervals. Those who are found to be eligible participants will be invited to answer some additional questions about mood and mood-related behaviors, and will receive incentives for doing so (see Tables 1 and 2 in Chapter 4). If they agree to participate in the trial of daily versus monthly online monitoring of mood, they will be assigned at random to one these two conditions.

Some participants may opt for an email, telephone, or mailed reminder about when to log in (e.g., a daily reminder for those assigned at random to the daily monitoring condition; a monthly reminder for all others). The study team has devised a method that allows these participants to contact a third party vendor, and to provide an email address, telephone number, or street address or PO Box to which the reminder can be delivered. This reminder operation occurs after the participant has logged out of the LSE system, and the research team will not

have access to these potentially identifying facets of personal information (email, telephone number, mailing address). This login reminder has been used in other studies with the LSE. This will be assigned as a non-randomly assigned covariant in the statistical analysis for participants who identified they would use such a reminder service.

Several protections will be taken to ensure the confidentiality of study participants, which will be safeguarded in the following ways: (1) use of coded subject identification numbers, (2) the investigative team will use code number data only for analysis; (3) no identifying information, email addresses, and/other potential identifiers will be collected by the research team. Although data should be entirely electronic, any printed data will be stored in a locked file cabinet when not in use and will not be accessible to personnel other than the investigators or data collectors. Data will be stored in secure, password protected computer files located in the MSU Communications Technology Lab and in the Department of Epidemiology at MSU.

SECTION 5.D: Potential risks

Research risks if any are no more than minimal. It is possible that some participants may find the online surveys to be an onerous task, even though the daily log-ins are designed to require no more than a few minutes of time each day.

The study team will not be gathering personally identifying information and the participants will be instructed to choose userid and passwords that cannot be

used to identify them as individuals, except to themselves. The contract with the third party vendor for email, telephone, and postal reminders will include a non-disclosure clause with tangible penalties if the vendor discloses or shares this information with others.

The only other tangible risk might be frustration, sad feelings, or the like, if the content of the survey items reminds the participant of unfortunate circumstances (e.g., persistence of smoking in the presence of a desire to quit). In many studies of this type, one foreseeable risk is breach of confidentiality due to unintended release of research data. However, in this study, the participation is anonymous and we are not proposing to gather any survey data that can be linked to individuals.

SECTION 5.E: Protection against risk

We have strived to streamline our assessments so as to minimize subject burden, intrusiveness, and inconvenience. As described earlier, this protocol has been designed to for subjects to participate anonymously. In addition, the protocol has implemented methods to direct subjects in crisis to mental services. The protocol also includes reminders to participants on how to access mental health services on campus under any circumstances.

SECTION 5.F: Potential benefits

In addition to the monetary reward for participation, subjects will benefit from the experience of participating in scientific research that may yield benefits

of public health significance. The nature of the public health benefit consists of the contribution to more effective assessment of mood, problem drinking, or drug use through the use of anonymous Internet surveys.

Participants may benefit by learning more about their own health habits, he/she may seek assistance if there are difficulties resulting from depression, substance or tobacco use. Monitoring of mood use may encourage some students to be more aware of their own mood state and its impact on their lifestyle, which would be a benefit

SECTION 5.G: Inclusion of Women

It is anticipated that female as well as male college students will participate in the survey. However, it is not the intent of this study to specifically recruit either males or females into the study. The study will seek participation of women at a rate proportional to population of interest by selection of recruitment sites (dormitories) whose residence gender distribution is consistent with the University population.

SECTION 5.H: Inclusion of Children

It is anticipated that participants under the age of 21 will make up the majority of those recruited due to the population of interest in this study. To the extent that children are defined to include individuals who have not yet reached the age of majority (less than age 18 years) and legal status to sign their own consent forms, children will be included. However, for college-attending youths

who are in this status (e.g., who have entered as freshmen at age 16-17 years), we will arrange for their participation by having the third party vendor provide a consent form that can be signed by a parent and returned to the vendor. The PI will inspect but will not retain a copy of the parent's signed consent, which otherwise could serve as a linking identifier, and if desired by the IRB, a copy of the signed consent form will be filed with the IRB. Of course, because the surveys are conducted anonymously, we will have no way to verify the age of the respondent.

SECTION 5.I: Inclusion of Minorities

It is anticipated that minorities will participate in the study. However, it is not the intent of this study to specifically recruit minorities into the study. The study will seek participation of minorities at a rate proportional to population of interest by selection of recruitment sites (dormitories) whose residence diversity distribution is consistent with the University population.

SECTION 5.J: Identification of adverse effects

Subjects will be provided instructions for contacting either the Principal Investigator for the study, or the chair of the MSU human subjects committee should he/she have any reports of adverse events, although we envision none. When participants are residents of official university residences (e.g., dormitories), the residential hall directors will be informed of the study, and will be given instructions to follow if and when they become aware of adverse events, or

if they or others have questions or concerns. Given the nature of this study, we have not specified a Data Safety and Monitoring Board, except to have the PI and Co-PI serve the functions of the DSMB, as is customary in no more than minimal risk trials of this nature, but in addition a small DSMB will provide an oversight function.

SECTION 5.K: Data and Safety Monitoring Plan

The PI and Co-PI will implement a data and safety monitoring plan, and will use the LSE to maintain the database elements should adverse events or other data safety issues arise.

SECTION 5.L: Data Safety and Monitoring Board

LSE provides a sophisticated data entry module, which interacts with the analysis capabilities of Excel; Thereby, the PI and Co-PI will have immediate access to data analysis results as soon as data are entered. Quality checks will be assessed at three points: RAs functioning as data entry personnel will run initial frequency checks, followed by proofing, and a final proof on key outcomes by the project manager and LSE staff. The principal investigator and Co-PI will be provided password-protected access to the data files, along with codebooks on all forms/data files, as will the chairman of the MSU IRB if he wishes to inspect the DSM files. No personal information such as name or social security number will be entered into any of the SPSS databases. If there is need to retain personal information or contact information (e.g., a complaint about the study or

the behavior of an RA or other member of the project staff), this information will be forwarded to the IRB for storage; it will not be kept by the PI or Co-PI except via this IRB-mediated mechanism.

For oversight of this data collection, transfer, entry, and plan for review, we will organize a Data Safety and Monitoring Board, which will meet at least once yearly during the course of the project. Additional meetings will be organized on an as-needed basis. This board will consist of a subset of the MSU IRB members (current or former).

SECTION 5.M: Quality Assurance

Quality assurance procedures are built into the automated LSE procedures and the processes of recruitment. These procedures will be audited and spot-checked on a periodic unannounced sample basis by the PI or her designee.

REFERENCES

Allen, J.P., Litten, R.Z., Fertig, J.B., & Babor, T.F. (1997). A review of research on the Alcohol Use Disorders Identification Test (AUDIT). *Alcoholism: Clinical and Experimental Research, 21,* 613-619.

Armstrong, J. (1975). Monetary incentives in mail surveys. *Public Opinion Quarterly, 39,* 111-116.

Angrist JD, Imbens GW, Rubin DB. (1996). Identification of causal effects using instrumental variables (with Discussion). *J Am Statistical Assoc* **91**:444-472.

Anthony, J.C., Neumark, Y., Van Etten, M.L.. Do I Do What I Say? (2000). A Perspective On Self Report Methods In Drug Dependence Epidemiology. In A.A. Stone et al., Eds. *The Science of Self Report.* Lawrence C. Erlbaum Publishers. Mahwah, New Jersey.

Babor, T.F., Biddle-Higgins, J.C., Saunders, J.F., & Monteiro, M.G. (2001). AUDIT: *The alcohol use disorders identification test: guidelines for use in primary health care*. Geneva: World Health Organization.

Barnard, J., Frangakis, C.E., Hill, J.L., & Rubin, D.B. (2003). A principal stratification approach to broken randomized experiments: A case study of school choice vouchers in New York City. *Journal of the American Statistical Association (with discussion), 98,* 299-323. Full publication at http://www.biostat.jhsph.edu/~cfrangak/papers/sc/vouchers.pdf

Birnholtz, J.P., Horn, D.B., Finholt, T.A., & Joo-Bae, S. (2004). The effects of cash, electronic, and paper gift certificates as respondent incentive for a web-based survey of technologically sophisticated respondents. *Social Science Computer Review,* 22(3), 355-362.

Bosnjak, M., & Tuten, T.L. (2003). Prepaid and promised incentives in web surveys. *Social Science Computer Review,* 21(2), 208-217.

Campbell, D.T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 54, 297-312.

Church, A.H. (1993). Estimating the effect of incentives on mail survey response rates: a meta-analysis. *Public Opinion Quarterly, 57,* 62-79.

Cochran, WG. (1968). The effectiveness of adjustment by subclassification in

removing bias in observational studies. *Biometrics* **14**:295-313.

Cole, J.C., Rabin, A.S., Smith, T.L., & Kaufman, A.S. (2004). Development and validation of a Rasch-derived CES-D short form. *Psychological Assessment, 16(4),* 360-72.

Denicoff KD, Smith-Jackson EE, Disney ER, Suddath RL, Leverich GS, Post RM. (1997). Preliminary evidence of the reliability and validity of the prospective life-chart methodology (LCM-p). *J Psychiatr Res.* Sep-Oct;31(5):593-603.

Denicoff KD, Leverich GS, Nolen WA, Rush AJ, McElroy SL, Keck PE, Suppes T, Altshuler LL, Kupka R, Frye MA, Hatef J, Brotman MA, Post RM. (2000). Validation of the prospective NIMH-Life Chart Method (NIMH-LCM-p) for longitudinal assessment of bipolar illness. *Psychol Med.* Nov;30(6):1391-7.

Denicoff KD, Ali SO, Sollinger AB, Smith-Jackson EE, Leverich GS, Post RM. (2002). Utility of the daily prospective National Institute of Mental Health Life-Chart Method (NIMH-LCM-p) ratings in clinical trials of bipolar disorder. *Depress Anxiety.* 15(1):1-9.

Daley, E.M., McDermott, R.J., McCormack-Brown, K.R., & Kittleson, M.J. (2003). Conducting web-based survey research: a lesson in Internet designs. *American Journal of Health Behavior, 27(2),* 116-24.

Downes-Le Guin, T., Janowitz, P., Stone, R., & Khorram, S. (2002). Use of pre-incentives in an Internet survey. Journal of Online Research. http://www.ijor.org/archives/articles/use_of_pre-incentives_in_an_Internet_survey.pdf

Frangakis, C.E., & Rubin, D.B. (1999). Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika, 86,*(2), 365-379. http://biomet.oxfordjournals.org/cgi/reprint/86/2/365

Frangakis, C.E. & Rubin, D.B. (2002). Principal stratification in causal inference. *Biometrics, 58,* 21-29. http://www.biostat.jhsph.edu/~cfrangak/papers/preffects.pdf

Frangakis, C.E., Brookmeyer, R.S., Varadhan, R., Mahboobeh, S., Vlahov, D., & Strathdee, S.A. (2004). Methodology for evaluating a partially controlled longitudinal treatment using principal stratification, with application to a needle exchange program. *Journal of the American Statistical Association,99(465), 239-249.* *http://repositories.cdlib.org/cgi/viewcontent.cgi?article=4084&context=postprints*

Helzer, J.E., Badger, G.J., Rose, G.L., Mongeon, J.A., & Searles, J.S. (2002).

Decline in alcohol consumption during two years of daily reporting. *Journal of Studies on Alcohol, 63,* 551-558.

Heerwegh, D. (2006). An Investigation of the Effect of Lotteries on Web Survey Response Rates. *Field Methods* 2006; 18; 205.

http://www.census.gov/prod/2005pubs/p20-554.pdf

http://www.epi.msu.edu/janthony

http://www.nimh.nih.gov/publicat/students.cfm

http://moodGYM.anu.edu.au

http://www.pewInternet.org

Jo B. (2002). Model misspecification sensitivity analysis in estimating causal effects of interventions with non-compliance. Stat Med. Nov 15;21(21):3161-81.

Johnston, L.D., O'Malley, P.M., Bachman, J.G., & Schulenberg, J.E. (2004). *Monitoring the future national survey results on drug use, 1975-2003. Volume II: College students and adults ages 19-45* (NIH Publication No. 04-5508). Bethesda: National Institute on Drug Abuse, 267 pp.

Kephart, W.M., & Bressler, M. (1958). Increasing the response to mail questionnaires: A research study. *Public Opinion Quarterly,* 22(2), 123-32.

Kisch J, Leino EV, Silverman MM. (2005). Aspects of suicidal behavior, depression, and treatment in college students: results from the spring 2000 national college health assessment survey. *Suicide Life Threat Behav.* Feb;35(1):3-13.

Leverich GS, Post RM: *NIMH LC Manual for Recurrent Affective Illness.* Bethesda, National Institute of Mental Health, 1993.

Leverich, G. S. & Post, R. M. (1997a). *The NIMH Life Chart Manual for Recurrent Affective Illness : The NIMH-LCM™.* Biological Psychiatry Branch Monograph, Revised. National Institute of Mental Health: Bethesda, MD. http://www.bipolarnews.org/pdfs/Patient%20Prospective%20Manual.pdf

Leverich, G. S. & Post, R. M. (1997b). *The NIMH Life Chart Manual-Self Version™: The NIMH-LCM-S.* Biological Psychiatry Branch Monograph, Revised. National Institute of Mental Health: Bethesda, MD. http://www.bipolarnews.org/pdfs/Patient%20Retrospective%20Manual.pdf

Li F, Frangakis CE. (2005). Designs in partially controlled studies: messages from a review. *Stat Methods Med Res.* Aug;14(4):417-31.

Li F, Frangakis CE. (2006). Polydesigns and causal inference. *Biometrics.* Jun;62(2):343-51.

Lieberman JA, Greenhouse J, Hamer RM, Krishnan KR, Nemeroff CB, Sheehan DV, Thase ME, Keller MB. (2005). Comparing the effects of antidepressants: consensus guidelines for evaluating quantitative reviews of antidepressant efficacy. Neuropsychopharmacology. Mar;30(3):445-60.

March J, Silva S, Petrycki S, Curry J, Wells K, Fairbank J, Burns B, Domino M, McNulty S, Vitiello B, Severe J (2004). Treatment for Adolescents With Depression Study (TADS) Team. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents With Depression Study (TADS) randomized controlled trial. *JAMA.* Aug 18;292(7):807-20.

Mock V, Frangakis C, Davidson NE, Ropka ME, Pickett M, Poniatowski B, Stewart KJ, Cameron L, Zawacki K, Podewils LJ, Cohen G, McCorkle R. (2005). Exercise manages fatigue during breast cancer treatment: a randomized controlled trial. Psychooncology. Jun;14(6):464-77.

Morgan, C., & Cotton, S.R. (2003). The relationship between Internet activities and depressive symptoms in a sample of college freshmen. *Cyberpscyhological Behavior, 6*(2), 133042.

Pealer, L.N., Weiler, R.M., Pigg, R.M., Miller, D.M., & Dorman, S.M. (2001). The feasibility of a web-based surveillance system to collect health behavior data from college students. *Health Education and Behavior,* 28(5), 547-59.

Pealer, L., & Weiler, R.M. (2003). Guidelines for designing a web-delivered college health risk behavior survey: lessons learned from the university of Florida health behavior survey. *Health Promotion Practice,* 4(2), 171-9.

Radloff, L.S. (1977). The CED-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 384-401.

Rosenbaum P, Rubin DB. (1984). Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc* **79**:516-524.

Santor DA, Poulin C, LeBlanc JC, Kusumakar V. (2007). The use of monetary incentives in a community survey: Impact on response rates, data quality, and cost. *J Am Acad Child Adolesc Psychiatry.* Jan;46(1):50-9.

Saunders, J.B., Aasland, O.G., Babor, T.F., de la Fuente, J.R., & Grant, M. (1993). Development of the alcohol use disorders screening test (AUDIT). WHO collaborative project on early detection of persons with harmful alcohol consumption, II. *Addiction, 88,* 791-804.

Scholder, T., Barton, G., Gearan, P., & Casey, P. (2001). *The effects of data collection method and monetary incentive on response rates and data integrity in studies of prospective college students.* Denver: Paper presented at the 12th Annual Symposium. Maguire Associates, Inc.

Searles, J.S., Helzer, J.E., & Walter, D.E. (2000). Comparison of drinking patterns measured by daily reports and timeline follow back. *Psychology of Addictive Behaviors, 14(3),* 277-286.

Searles, J.S., Helzer, J.E., Rose, G.L., & Badger, G.J. (2002). Concurrent and retrospective reports of alcohol consumption across 30, 90, and 366 days: Interactive voice response compared with timeline follow back. *Journal of Studies on Alcohol, 63,* 352-362.

Shaw, M.J. (2001). The use of monetary incentives in a community survey: Impact on response rates, data quality, and cost. *Health Services Research, 35(6),* 1339-1346.

Storr, C.L., Reboussin, B.A., Anthony, J.C. (2004). Early childhood misbehavior and the estimated risk of becoming tobacco dependent. *American Journal of Epidemiology, 160(2),* 126-130.

The American College of Health Association. The American College Health Association National College Health Assessment (ACHA-NCHA), Spring 2003 Reference Group report. J Am Coll Health. 2005 Mar-Apr;53(5):199-210.

Tuten, L.T., Galesic, M., & Bosnjak, M. (2004). Effects of immediate versus delayed notification of prize draw results on response behavior in web surveys: an experiment. *Social Science Computer Review, 22(3),* 377-384.

Volk, R.J., Steinbauer, J.R., Cantor, S.B., & Holzer, C.E. (1997). The Alcohol Use Disorders Identification Test (AUDIT) as a screen for at-risk drinking in primary care patients of different racial/ethnic backgrounds. *Addiction, 92(2),* 197-206.

Wight, R.G., Sepulveda, J.E., & Aneshensel, C.S. (2004). Depressive symptoms: How do adolescents compare with adults? *Journal of Adolescent Health.* 34(4), 314-23.

Wong JG, Cheung EP, Chan KK, Ma KK, Tang SW. (2006) Web-based survey of depression, anxiety and stress in first-year tertiary education students in Hong

Kong. *Aust N Z J Psychiatry*. Sep;40(9):777-82.

Wu, L.T. & Anthony, J.C. (1999). Tobacco smoking and depressed mood in late childhood and early adolescence. *American Journal of Public Health, 89(12),* 1837-40.

Wu, L.T. & Anthony, J.C. (2000). The use of the case-crossover design in studying illicit drug use. *Substance Use and Misuse, 35(6-8),* 1035-50.